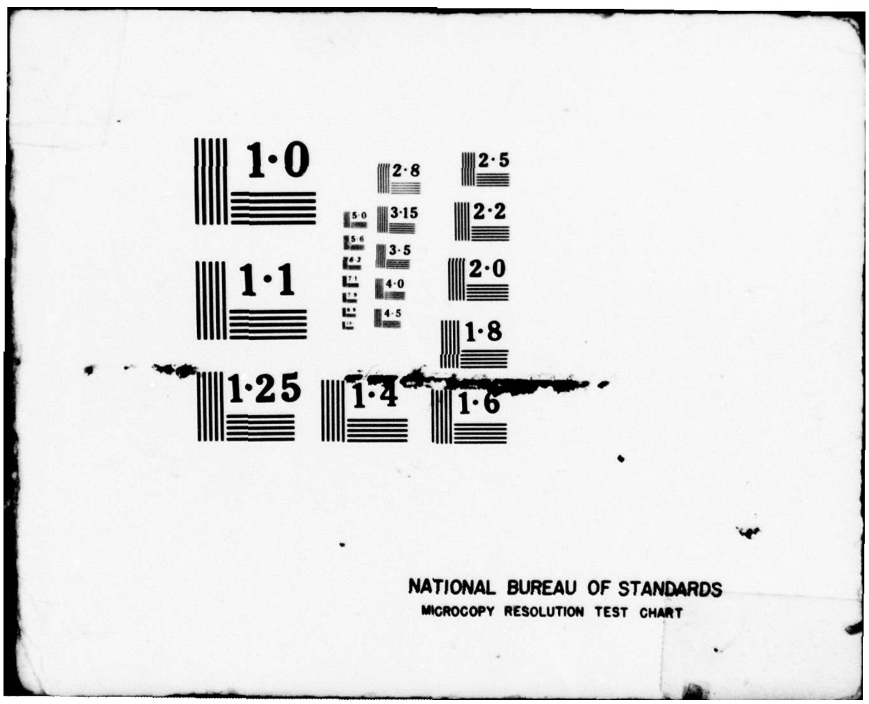
AD-A068 221 AIR FORCE ACADEMY COLO F/G 5/10
PROCEEDINGS OF THE ANNUAL SYMPOSIUM ON PSYCHOLOGY IN THE DEPART--ETC(U)
APR 78 J M BERMUDEZ, W H CLOVER, L J DAHLE
UNCLASSIFIED USAFA-TR-76-12 NL

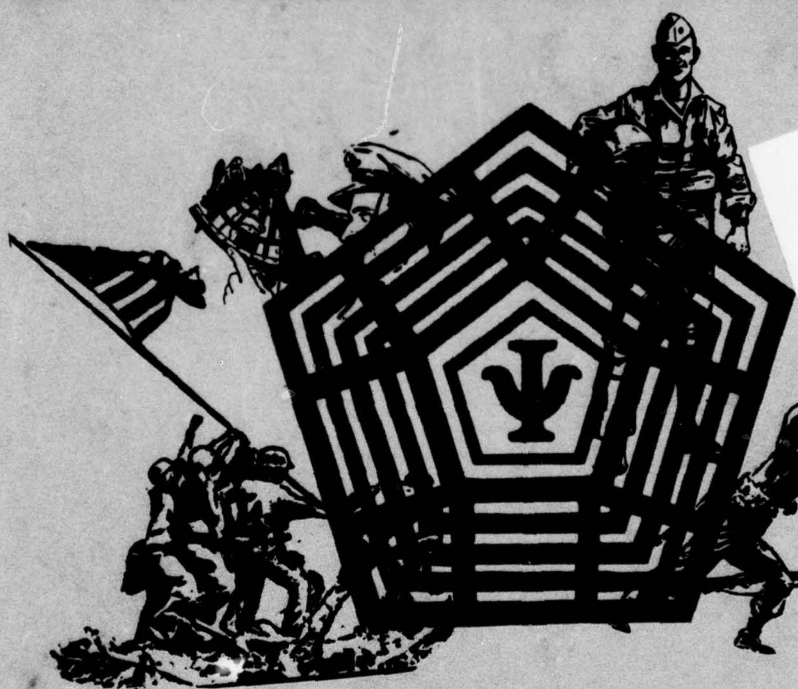
1 OF 4
ADA
068221


1·0
2·8
2·5
5·0
5·6
3·15
2·2
3·5
2·0
1·1
4·0
4·5
1·8
1·25
1·4
1·6
NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

LEVEL II

PROCEEDINGS

PSYCHOLOGY IN THE DEPARTMENT OF DEFENSE



SIXTH SYMPOSIUM

APRIL 1978

SPONSOR:

USAF ACADEMY

DEPARTMENT OF BEHAVIORAL SCIENCES AND LEADERSHIP

3

PROCEEDINGS OF THE SIXTH SYMPOSIUM

PSYCHOLOGY IN THE DEPARTMENT OF DEFENSE

21 April - 22 April 1978

Department of Behavioral Sciences and Leadership

United States Air Force Academy

Colorado Springs, Colorado

Sypmosium Chairman

Major Gene A. Berry

EDITORS

John M. Bermudez, Major, USAF
Captain William H. Clover
Captain Lee J. Dahle
Captain Richard L. Hughes
Captain Donnell M. Washington

COMMITTEES FOR THE SIXTH SYMPOSIUM

ASSISTANT SYMPOSIUM CHAIR

Lieutenant Colonel William Rosenbach
Captain Bruce Smith

EXECUTIVE COMMITTEE

Major Gene Berry, Symposium Chair
Lieutenant Colonel Lloyd R. Chason
Lieutenant Colonel Lawrence F. Sharp
Lieutenant Colonel Jock C.H. Schwank
Lieutenant Colonel Eugene H. Galluscio

PROGRAM COMMITTEE

Major John M. Bermudez, Chair
Captain William H. Clover
Captain Lee J. Dahle
Captain Richard L. Hughes
Captain Donnell M. Washington

INVITATIONS AND ATTENDANCE COMMITTEE

Major Valentin W. Tirman, Jr., Chair
Major John M. Anderson
Captain Dickie A. Harris

SOCIAL, BILLETING & TRANSPORTATION COMMITTEE

Captain B. Don Sullivan, Chair
Major Robert B. Linden
Captain Ronald LaScala
Captain Dean H. Wilson
Captain David C. Gillman

FACILITIES AND VISUAL AIDS COMMITTEE

Captain Jack D. Hannig, Chair
Captain Larry L. Wheeler
TSgt Frank C. Derry
TSgt Dale E. Schimmel

PANEL DISCUSSION COMMITTEE

Lieutenant Colonel Jefferson M. Koonce, Chair
Captain Blaine N. Lee, Jr.
Captain Thomas J. Twardowski

PROTOCOL COMMITTEE

Captain Robert W. Nordeman, Chair
Captain William L. Derrick

This research report is presented as a competent treatment of the subject, worthy of publication The United States Air Force Academy vouches for the quality of the research, without necessarily endorsing the opinions and conclusions of the author.

This report has been cleared for open publication and or public release by the appropriate Office of Information in accordance with AFR 190-17 and DODD 5230.9. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service.

This research report has been reviewed and is approved for publication.

John P. Wittry

JOHN P. WITTRY, Colonel, USAF
Vice Dean of the Faculty

Additional copies of this document are available through the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22151

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: (14) USAFA-TR-76-12
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): Proceedings, Psychology in the Department of Defense, 6th Annual Symposium
5. TYPE OF REPORT & PERIOD COVERED: Proceedings
6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): (10) [redacted] John M. Bermudez, [redacted] William H. Clover, [redacted] Lee J. Dahle, [redacted] Richard L. Hughes, [redacted] Donnell M. Washington
8. CONTRACT OR GRANT NUMBER(s): (12) 344p.
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Behavioral Sciences and Leadership
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: (11) Apr 78
11. CONTROLLING OFFICE NAME AND ADDRESS: DBL, USAF Academy, CO 80840
12. REPORT DATE: April 1978
13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report): Approved for Public Release; Distribution Unlimited

(6) Proceedings of the Annual Symposium on Psychology in the Department of Defense (6th), 21 [redacted]-22 April 1978,

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)
18. SUPPLEMENTARY NOTES
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Psychology	Pilot Performance	Attrition	Instructional System Development
Symposium	Organizational Development	Stress Management	
Personnel	Job Enrichment	Human Factors	
Training	Women in the Military	Simulation	
Training Research Evaluation	Visual Display and Processing	Leadership	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

These printed proceedings included papers and presentations that deal with a wide range of psychology and research on behavior: Personnel, Women, Training, Human Factors Engineering, Organizational Development, Simulation, Performance, Leadership, Attrition, and Stress.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

011 550 JOB

FORWARD

The opinions expressed in this volume are those of the individual authors and do not necessarily reflect official policy either of the Army, Navy, Air Force or of any government organization in which the author may be serving.

Trade names of materials or products of commercial or non-government organizations are cited only where essential to precision in describing research procedures or evaluation of results. Their use does not constitute official indorsement or approval of the use of such commercial hardware or software.

These proceedings are approved for public release, distribution unlimited, in accordance with paragraph IVA, AFR 80-45.

In publishing the Proceedings of the Sixth Symposium, the editors sought to facilitate an effective and timely dissemination of the technical information presented. The papers contained in this document were copied directly from unedited reproducible copy submitted by the authors who are solely responsible for their contents.

Acknowledgements

Cover and Layout, Graphics, Division of Instructional Technology,
Dean of Faculty, USAF Academy

Editorial Assistance: Janice Bishop

OVERVIEW OF SYMPOSIUM SCHEDULE

THURSDAY 20 APRIL 1978

Morning Activities

Event

Registration

Keynote Address
Dr. Theodore H. Blau

Paper Sessions

Organizational Development/Job Enrichment

Training and Training Research

Evaluation

Invited Luncheon Address
Dr. William L. Lehmann

Afternoon Activities

Panel Discussion & Paper Session

Panel: Organizational Development in the DoD: Process or Structure?

Human Factors

Paper Sessions

Leadership

Organizational Development/Job Enrichment

Human Factors

FRIDAY 21 APRIL

Morning Activities

Paper Sessions

Leadership

Women in the Military

Panel Discussion & Paper Session

Panel: A Multidisciplinary View of Leadership in Institutional - Occupational Framework

Training and Training Research

Afternoon Activities

Panel Discussion and Paper Session

Panel: The Expanding Role of Women in the Military

Human Factors/Simulation

Paper Sessions

Stress and Stress Management

Evaluation

Personnel

SATURDAY 22 April

Morning Activities

Paper Sessions

Human Factors

Personnel

Paper Sessions

Human Factors

Training and Curriculum

END OF SYMPOSIUM

PROCEEDINGS OF THE SIXTH SYMPOSIUM ON PSYCHOLOGY IN THE DEPARTMENT OF DEFENSE

THURSDAY, 20 APRIL 1978

FRIDAY, 21 APRIL 1978

SATURDAY, 22 APRIL 1978

PAPER SESSION: Personnel
Chair: Captain Dean H. Wilson, Department of Behavioral Sciences and Leadership, USAF Academy, CO.

The methodology and results of a research project designed to evaluate the benefits of conducting a special motivation program for first-term airmen within Military Airlift Command. Results from test and control bases are compared, using administrative and judicial, as well as survey data, as criteria.

PAPER SESSION: Human Factors
Chair: Captain Donnell M. Washington, Department of Behavioral Sciences and Leadership, USAF Academy, CO.

PAPER SESSION: Training and Curriculum
Chair: Captain Larry L. Wheeler, Department of Behavioral Sciences and Leadership, USAF Academy, CO.

KEYNOTE ADDRESS

by

Theodore H. Blau

QUO VADIS BEHAVIORAL SCIENCE

Colonel Williams, I see many old friends, colleagues, I hope some new colleagues, and participants in the Sixth Biannual Symposium. First let me thank you for inviting me. I'm really pleased to be here. I was graciously invited to the Fifth Symposium two years ago and was not able to attend. After reading the proceedings, I think that it turned out to be a good idea. They were excellent: undimmed and unspoiled by any kind of diverse presentations or keynote addresses. I suppose in the past three years, I've given upwards of fifty keynote addresses. Now the safest keynote address that you can give is Biblical: "Repent, repent you've been following golden idols, but it's not too late. Repent, there is a truer course of action." So, I suppose my keynote address has some element of this in it. I'll try to keep it down to a reasonable, sort of background signal and a minimum of noise.

I was asked to speak about the future of Behavioral Sciences in the Department of Defense. I enjoy speaking about that, because I've been associated with the Department of Defense, before it was the Department of Defense. I was telling some of our colleagues at dinner last night, I was getting rather depressed at looking at ribbons, I don't see the EPO Ribbon and the Victory Medal of World War II, which was my era. And I was telling our colleagues that on May 13, it will be the 31st anniversary of my discharge from the United States Army Air Corps, which was commonly known as the Old Brown Shoe Air Corps. So, in spite of the fact that you've made me feel very old, I'm really very glad to be among colleagues.

The past couple of years, I've worked for the Surgeon General of the Air Force, but more recently with Combat Arms at Fort Benning and with TRADOC with the United States Army as well as with the Surgeon General of the United States Army. So, I'm kind of familiar with the pitfalls, the problems, the opportunities, and the excitements that one faces almost regularly being part of the DOD.

I think if we were going to talk about the next decade, we should talk about 1980 to 1990 rather than 1978 to 1988. Behavioral Sciences in the Department of Defense for the decade 1980 to 1990 is, with very little question, going to be a reflection of the progress or lack of progress of the behavioral sciences in the overall general American comuunity. I say the overall, general American community, simply because, the largest segment of behavioral science research and application activity in the world occurs in the United States, unquestionably and without a doubt.

The Department of Defense and its predecessors have long been some of the most significant supporters of behavioral sciences research and application. It's the degree of criticism and focal vituperativeness that will fall upon the Armed Services from people who should know better in the academic and general community: The Armed Services are this and the Armed Services are that. During the past year, I've had the pleasure of pointing out to many of our colleagues in Washington and the APA and other learned societies, that it's the United States Armed Services where equality of opportunity was first established, where racial barriers were first lowered, and the usual derogatory, negative attitudes about the service are simply unsubstantiated in many respects. It's very frustrating to be part of a research and application sector that does a lot of the work and also gets very little of the accolade and much of the criticism; however, it's all in a good cause.

The kinds of things that are going to happen in behavioral sciences, 1980 to 1990 in the DOD, are going to be a reflection of how we end the 1970s. That's why I said, "Let's exclude 1978 to 1980," because the next Congressional session is going to be a very critical one for behavioral sciences. And if we end the 1970's well, you can look for 1980 to 1990 to be perhaps the most significant decade of support, activity, and productivity in behavioral sciences. And a good segment of that is very, very likely to occur in the government sponsored research and application in the Armed Forces. It is moving in that direction, although not as rapidly as many would like.

What is probably even more important, although this is certainly a very personalized opinion, is that the way we will see psychology and behavioral sciences in the 80's will be a reflection of the degree to which we get our act together. We are past the point where we can talk about experimental psychology, applied psychology, research psychology, clinical psychology, learning technology, clinical application, psychotherapy, and so forth, as though there are the good guys and the bad guys and the good gals and the bad gals. The Corpus Psychology is 64,000 members of APA, the American Psychological Association. So we are talking about a corps of over 80,000 well-trained, very bright human beings. And we are increasing our larder about 3,300 per year at the doctoral level. Psychology is the most, single most, popular elected undergraduate course in the United States and has been for seven years. Psychology is reaching an apogee of interest, availability, and knowledgeability in the general community. And, if we miss this opportunity to do all that psychology can indeed do to promote human welfare, we deserve the fate of the other learned professions who have not taken opportunities on the tide and run with it appropriately and we will deserve to sort of die out in the 1990s. But, it's in the 1980s that the story will to told.

There are a number of things that will be happening in the next two years that will determine the fate of the behavioral sciences in the DOD or in the general scientific and professional community. Three of them we have very little control about, and four or five we can do a great deal and are doing a great deal about, and I would certainly like to talk about this a little bit. The things that we're not going to be able to do much about, but which will have very significant influence, on the work that we're going to be doing, include the political climate in the United States, as well as the world; the economy in the United States; and world confrontation.

Now we here, even though we are involved in some of the things, are not going to be able to, very directly, oh well, some of the graduates of the Academy may well have some important role in that, but as behavioral scientists and psychologist, quite unlikely. On the other hand, there are things going on and happening now in places such as the National Institute of Mental Health, the National Science Foundation, of life, the question of training, professional development, and continuing education, and the question of our own consumers, the people who take our product and either find it useful or useless or confusing. These are the areas that can be influenced. Let me speak a little bit about both segments.

The political climate is a confusing one and has been for about 18 months. Our President is a different kind of President. Someone in Washington said: "Mr. Carter is a cross between a tiger and a canary. We don't know what he is, but when he whistles, by George, they listen."

I had the pleasure of meeting Mr. Carter, personally, last year when he signed the Mental Health Commission Bill. He invited a relatively small number of people to the White House to be present, and then afterwards got rid of the press, we went into the West Room, and

we got a chance to talk. My oldest son, who just graduated law school, heard I was going to the White House, and he called up and he said: You're not going to the White House in that old blue suit, with stretch fabric and wide lapels, that you got at the Sears Men's Store." I said, "Well, I'm not going to get a new suit just to go to the White House." He said, "You're going to embarrass us all." And he's a rather elegant, well-dressed guy. He says, "Why don't you go out and buy a new suit." I said, "I don't have your advantages. I don't have an affluent father." So I went to the White House in my two-button, stretch fabric, Sears Men's Store, blue suit and a light blue shirt and a blue tie with red dots. And as I was sitting in the East Room waiting for President and Mrs. Carter to come in, I began to think maybe Jeff was right, you know, I'll be embarrassed here, this hick from Tampa, Florida, wearing this out-of-date Sears Men's Store suit. And then suddenly the Commander, who is the liason officer, the tall submariner, who is Mr. Carter's personal military escort, came in and said: "Ladies and gentlemen, the President of the United States." We all stood up, and here comes Mr. Carter wearing a two-button, stretch fabric, dark blue suit and you know for the rest—I was there two and one-half hours—I could just barely control myself from saying: "Mr. President, may I see the label?"

Let me tell you something about the White House, it's a very familial White House: there are dogs running around, Amy's dog wanders around, and a couple of the other staff people have dogs. There are children there, he encourages the staff to alternate bringing their kids to see what their parents do. And, I think the most positive thing I can tell you about the Presidency is that during the two and one-half hours, Mr. and Mrs. Carter almost never took their hands off each other. They were patting, he had his arm around her, very warm, friendly, very natural couple. One doesn't see much of that in Washington. Usually the wife is sort of a nervous person, that follows around, worrying about doing something wrong, but President and Mrs. Carter are real home folks. They're very warm, very friendly. Mr. Carter's very positive about the possible contribution of the behavioral sciences. He and Mrs. Carter talk about the psychological health of the country. Their language, their attitude reflects a lot of positive feelings about psychology. They have both studied psychology, and I'm going to get back to that in a moment, because it makes a great difference in the corridors of power as to whether the people who make the decisions have had any exposure to what we consider our basic science of human behavior. But, like in all political matters, it's much more lip service than productivity.

Politics is the art of compromise and compromise is what you do when pressure comes from diverse sources. And psychologists and behavioral sciences have simply never had, up until two years ago, any kind of clout. But I think the time is coming where we're going to see some differences, which I'll tell you about. But, at the moment the political climate is relatively positive to behavioral science, but very subject to other kinds of direct pressures. In other words, from the executive branch of the government, you can expect a very interested group of people in respect to psychology and behavioral sciences, but you can also expect a lot of in-fighters, who have a lot of experience in Washington saying, "You don't want to give funds to these people. What the hell are they going to do for you? And how are they going to use the money? And it doesn't count anyway."

In the President's Commission on Mental Health, we have some representation in psychology. There's a lot of medicine, a lot of consumerism, but there are four or five of our best behavioral scientists heading task forces and they're doing their home work: the one on Crime Rape Prevention headed by George Alby, the one on cost effectiveness, John Conger's work, he's a Colorado fellow here at the Medical School, he's a member of the Commission, telling the people in Washington, that if you want to see it done right, if you want to see it done on data, if you want to see it done with some concern for what Mr. Carter likes to call the underserved and the needful of America. Psychologist, behavioral scientists, they're the people who are really thick into this and not politicians. So, on the political climate, there are some things going on that are somewhat positive, but we have no direct influence. The influence we're going to have there is good work, which will sort of come to the surface, if indeed it is good.

The most important thing that is coming up for all of us, even though most of you will disagree, is National Health Insurance, not because it is particularly going to affect you, but because in National Health Insurance, psychologists will be identified as available to the general public to meet their needs. If the record turns out to be outstanding, all of psychology will profit. If we bomb out, as psychiatrists have been traditionally bombing out, and promising much and delivering little, in terms of cost-benefit, cost effectiveness, every one of us is going to suffer. Kid yourself not. If you're in a research laboratory doing basic physiological research and a group of psychologists gets in the paper for being Nationl Health Insurance scams, cheating the public, doing bad work, your name as a psychologist is going to be smirched and your going to be hurt, and research funds are going to be harder to get. If on the other hand, the Colorado studies on Medicare, which are now going on, the CHAMPUS studies which APA has been commissioned to do on the effectiveness of the delivery of psychological services, turn out to be outstanding, everybody is going to find that the next time we talk basic research funds, the next time we talk training funds, the ears are going to listen better and the responses are going to be more positive. It's in with the touches and in with the falls. Whoever does good in behavioral science affects their colleagues and whoever does bad, affects them very strongly.

The economy has tremendous impact on what we all do, as you well know. We've been living in kind of a desert for a while as a result of some of the things that have occurred in the 1970s in the way of the recession and in respect to the government, the Office of Management and Budget's response to what to do with the available funding sources. Behavioral science has always been considered a second or third priority. This has started to be questioned in a number of ways: Maybe we really need these people? Maybe they have some answers to the insloube problems? Suddenly, when a politician begins to think: "one of the real problems is that people don't like their work and so they don't do good work. And it costs more to keep them doing bad work. And people come into the Air Force and then six months later they take an early discharge, they're not fit for the service, it costs us 18,000 dollars-Gee, you know, somebody ought to take a look at these problems. Who should it be?" And they're beginning to think about behavioral science and psychology. So, the economy, even though it has given us some difficulties in the past seven years in terms of research and training funds, has begun to prod those who walk the corridors of power to ask the question: "Since we haven't been able to improve the situation, are there some people who could help us?" And they are beginning to look toward us. But we have no direct capacity for input to the economy. Recently, we've had some positive indicators. I don't know what's going to happen. The economists are the only people who are foolish enough to make specific predictions about the economy. But, we're going to be victims or products of what happens in the economy in the next couple of years.

Some of our worst enemies could be our best friends. I heard today on the early television news, that Senator Proxmire gave another Golden Fleece Award to a behavioral scientist. The behavioral scientist had been interviewing prostitutes on a 93,000 dollar NSF grant. And Senator Proxmire said, "You know, this is sheer nonsense. Waste of money. Golden Fleece Award." The psychologist involved, damned fool that he is, issued a public statement that: "A Golden Fleece Award from Senator Proxmire is second only to a Pultizer Prize." Well, that's cute, but that's dumb, and that will hurt every one of us.

Senator Proxmire is a strange and interesting man. He's a jock, he runs, he exercises. From May 15 until November, he swims in the Washington Hilton Pool. It's a pool where you can go in at seven in the morning and there's no gates, and it's the only one in Washington. So, being that I am a professional swimmer, I always stay there. And I do my mile laps and he's lapping away. I met him about five years ago--I didn't know who he was--this funny guy with the checker-board head. You know, when his hair is wet, it's like a checker-board, is doing what I call the Minnesota crawl--like this. And he's very vigorous and I'm doing my laps and I'm puffing at the end and I lift my goggles and he's puffing right next to me and I said: "Say, I noticed that you're a vigorous swimmer and it happens to be my line. You know, if you were to stretch your stroke out and use the guideline of the marker in the pool and just reach for it and the midline, your stroke would smooth out, you'd go faster." Turns to me and interupts: "If I wanted your God damn advice, don't you think I'd ask for it." I said: "Well, yeh, of course." And, you know, he left, and then the pool boy came in. I said: "That's a pretty feisty guy, is he a regular?" He says: "That's Senator Proxmire." Oh.

So the next morning I come down, and he's chopping away with the Minnesota crawl, and the pool boy came in and he jumps out and says: "There's not enough chlorine in this pool, see that you get it up to snuff." "Yessir." And off he goes. And I think, boy what a feisty guy and suddenly I think: "I've had a pool for twelve years, you can tell when there's too much chlorine, but without a test kit, there's no way in hell you can tell when there's not enough." A very positive and definitive guy.

Well, over the months, we sort of got to talking a little friendlier. He asked me about my goggles. I said, "I have an extra pair, take them." Oh no. He's very sensitive about taking any kind of bribery, including a two dollar pair of swimming goggles. But I want to tell you this, at one time he gave a speech about psychology and psychiatry are no damn good because they sit in their offices in big cities and they never go out in the country and they're not available. And I scribbled a little note with the map of the United States, that we prepared (for) APA of the dispersion of psychologist in the United States, and it pretty well matrixes the 1970 census. And I said: Dear Senator Proxmire, If I can impose upon our swimming relationship, I'd like to call to your attention that you put psychology and psychiatry in the same boat geographically, and it's not quite so." And he read it into the record in the Senate about four days later. Now there's a moral to that story, and that is: powerful people are not entirely unreasonable. Senator Proxmire has gotten his reputation by being a tough, mean, aggresive, humorous, sarcastic person who says behavioral science ought to make sense. And it seems to me, that if we can find a way to convince him that a lot of what we do--and Lord know it's in the DOD--that a lot of what psychology and behavioral sciences does, makes sense: our work and organizational effectiveness; the primary leadership training that we're doing in the Army; the revision of the recruitment system. There's a lot of practical stuff and somebody ought to be bringing this to his powerful attention. Well the fact of the matter is, we're beginning to do that. But the economy can reverse itself in terms of its impact, in that people in the corridors of power say, "Show us something that makes sense." And behavioral science is getting to the point where they might be able to do that.

Well, confrontations are a tremendously important, growing area of concern that are going to affect everybodies lives. It already affects the economy and your life. You well know that you can no longer put your service pistol in your briefcase and carry it around with you, because when you go to the airport you have to go through a screening situation, for which you pay anywhere from 35 to 65 cents a person for every flight you take. Terrorism, hostage behavior, mass destructiveness, hijacking, these are phenomena of our era. There are probably less than 2,000 active people in that game, and yet thay can bring an entrie country to the brink of disaster. And they probably can do even more than they've done before. Well, we in behavioral sciences cannot directly influence that. We're not sure that anybody can directly influence it, but as some of you well know, in trying to cope with this, after they went through the usual procedures of using their best clandestine agency training, their best military, and police training, scientist to say: "Hey, can you help us with this. These are strange folks, that we don't understand, who create situations with which we're not equipped to cope." And they're very happy with the initial stuff that behavioral scientists have begun to produce. But we still don't have a very strong impact on this very significant and very dangerous, and very frightening aspect of modern life.

Well let me talk a moment about the things that I believe we can influence, in terms of what is going to happen to us in the years 1980 to 1990. The National Institute of Mental Health has been a medical facility since its inception. Psychologists have been very deeply involved, research psychologists, clinical psychologists, social psychologists, academics, a lot of funding has come through NIMH for our graduate schools. There is about a 60-40 chance that the next director of NIMH will be a psychologist. We have never been this close to such a role. And if this were to happen, we can rest assured that research funding will get a very, very careful review each and every year in terms of what psychologists and behavioral psychologists in general require, what they can potentially contribute, and how they should be supported.

Our first awareness of the kind of clout that we had with NIMH came just this year. We've continually fought against the cuts that have been occurring. For instance, let me tell you what happened in the 1979 preliminary budget, next year's funding. There was going to be a considerable cut in research training for psychologists. As a result of our activity in Washington, we not only reversed the cut, but gained a million and a half dollars for research training in 1979 and it's now tentatively budgeted at 15 million dollars. They plan to cut clinical training to the bone, literally, they simply said: "We'll train primary providers, medical students. They'll do all the mental health delivery in the primary prevention areas." We convinced them that it was impractical and we ended up with a 68 million dollar budget in clinical training. In the area of extramural research, which is where NIMH begins to support basic research, brain behavior research, physiological research, we not only reversed the cuts, but gained 21 million dollars for 1979. So extramural resaarch out of NIMH will go out at 106 million dollars. We were flabbergasted when this was the final budget figure. You know, we've worked hard, but we've worked hard for years and they say: "Yes we'll do our best. Yes. Fine." And in the end you find yourself at the bottom of the list. One of the exciting things was an increase in 17 percent in behavioral and neural sciences. NIMH will support research in behavioral and neural sciences at 33 million dollars next year. In the National Science Foundation, where we expected at best to hold our own, we are up 22 percent in research funding for the social sciences, 29 million dollars. As you know, the new director of the National Science Foundation, is a behavioral scientist, one of our outstanding ones, and it is very nice that we've come up from about 22 to 29 million dollars for the 1979 budget for research in social science.

This has been an outstanding product of the political activity of the American Psychological Association. Many of our academic colleagues, our research colleagues, say one of the things I don't like about the American Psychological Association is that they do so God damn much politicking. And it's an area where there's a deep concern, we're losing our academic colleagues because of this. Reality is, that most of the politicking as it's called, it's good staff work essentially, is in the direction of research funding and training funding, stuff that goes directly to our academic and research colleagues.

Personally, the next area is the one in my mind that will make our greatest impact in the next two years and that is in the area of consumerism. People have latched on to the Ralph Nader syndrome: What are you doing? How do you do it? How do we get our bucks worth? and Why do you lie to us? *The consumer movement is not growing, it's grown.* It's at an apogee. People want to know details. The people who are hurting are physicians. The patient comes in and starts out by saying: "What are you going to charge me for the examination? Do you base your fee on income? What is your success rate?" And they're flabbergasted. How dare a patient ask such questions.

Well, we have our patients in the DOD. They are people who are going to be responsible for the combat readiness of our troops, they're the training people, they're the leadership people. What are you doing? How much is it going to cost? What are we getting out of it? How applicable is it? They're even talking about things like: How much down the road evaluation do you have on this data? My God, they've gotten into our camp and started studying evaluation techniques. They're asking hard questions. It's embarrassing. But it's great, because of all of the areas of support in the DOD, behavioral science has one of the finest potentials for answering these questions in the future.

But it's true that we lie to them: lie by ommission. It's true that we hedge and we evade: Well, don't ask us too many questions, we don't know yet. It's true that we're willing to publish without admonitions that this must be replicated and we need some very solid, down the road evaluation to see if these techniques are really effective. It's true that we don't take the risk of doing measurable, behavioral responses in our evaluation research. Instead, we take a survey: How did you like the course? How did you like the training. Well, this is a wonderful course in organizational effectiveness, because 83 percent of the people that took the course, liked it. You could show a porno flick and you'd get a better percentage of positive response than that.

And so our friends, and they are our friends, who are asking hard questions--I was at Fort Levenworth last year at the Command General Staff School consulting with General Lewisell and he said: "I'm almost afraid to ask the question: But how in the hell do we know that this ONE is going to work?" I said, "That's the best question you've asked all day. Now if you'll support us in doing a down the road evaluation of the graduates 30, 90 days, 180 days after and 2 years after the training, compared with the match group that didn't have it, we're going to answer your questions." He says, "That's sensible." I said "I hope that we can get to the point where we can discuss sensible things and support them fiscally and then we're going to get some answers and we're going to have a much closer marriage than we've ever had before between DOD and behavioral scientists." So those people who are in the corridors of power, whether they be legislative people or military people, are becoming much more sophisticated about asking us the value of not only our applied research, but of our basic research and we have to get into a position to learn to stop looking offended, and feeling indignant that nonbehavioral scientists are asking us hard but realistic kinds of questions.

An area in which we can make tremendous input, which we haven't, is professional development. The match between our training and our function is poor. I am astonished at the magnificant results that come from our researchers, our applied researchers, our professionals, when you consider the gap between what they gave us in graduate school and the realities of the world. And how many times have you been asked by your professors at the graduate school: "Hey, how did we do?" "Five years later, did we give you anything worthwhile?" "Is there anything you have to tell us from the real world about what we should be doing with our students?" This is where we could have the most powerful input in the training that will produce those who are going to replace us in 20 and 25 years. Those of you out in the field have a tremendous amount of information and experience, which almost never goes back to the students. So they have to go out and spend 5, 10, and 15 years re-inventing the practical wheel and that's a crime. We are producing about 3,000 PhD's in psychology a year. Almost none of the young PhD's know a thing about what's going on in the real world and to some extent that's our fault. And if we could, in some way, establish a linkage with our institutional affiliations, with our friends the professors, the academics, with the heads of our old departments and say, "Look, I am a natural resource for behavioral science. If your people would only-even submit some grant requests for down-the-road-evaluation on the effectiveness of training in psychology, it would have a major impact in the style of training, both in research and in the professional areas of psychology." Now, how to do that, I don't know. But it is an auspicious, productive situation waiting for some clever people to establish the kind of linkages that can bring it about.

What else can we do? Well, maybe very little, but maybe a lot. One of the problems that we face, in my way of thinking, is severe, and intractable--and by intractale, I simply mean that I've spent about 3 years working on it, and I've had very little success. Other people tell me they feel the same way and that is: What can we do to get our act together? There's one group that says, I hope we never get our act together, and that the clinicians, and the experimentalists, and the physiologists continue to fight and argue, because this is a healthy kind of a base line setting in which new ideas can emerge. And I think there's much merit in that idea. On the other hand, psychologist and behavioral scientists, in general, would rather fight than cooperate. Particularly in the academic setting, argumentation, fighting, and dissension often results in very creative and seminal kinds of research. So, it's one of those deals where you got to be careful, you don't want too much and you don't want too little. But certainly where it comes to facing the people who are in a position to support us or reject us, this divided front is tremendously destructive.

Daniel Ionye, the Senator from Hawaii, is a very, very strong proponent of behavioral science. He's a modern thinking, tough, aggressive and, as you well know, a very powerful guy, as the Vice Chairman of the Armed Forces Finance Committee. Sitting in his office last fall, he said, "You know we get the worst dingbats that come in here and exert a lot of influence on funding for one reason alone: they got their act together and there's no dissension. We get outstanding professions like psychology coming in here and asking for something and then we start getting phone calls: Did you hear from so and so? Well, listen Senator, you don't want to go that route because that's not really good. And you got about eight different factions fighting with each other. So, you know what Senators and Congressmen do, they table it, they send it to a committee where it'll die, they can reassure everybody: "I'm working for you," but nothing ever happens. Politicians have got to learn that kind of technique, to say yes, but to end up with no. And the most positive, guaranteed, absolute way of going nowhere in the federal sector, is to have conflict within your own group. You in the military know this better than behavioral scientists who are not a part of the military. So in some way we have got to form a consortium of mutual interest, which is not too bad an idea, because we are moving to an era of time when the differences between clinical psychologists, and physiological psychologists, and social psychologists will be more superficial than real.

The interaction between the different kinds of psychology and behavioral sciences is intensifying, it's heating up, it's becoming very critical. One of the areas in which we see this happening is physiological psychology. For years we have been deeply concerned about the deceleration of training and job availability for well-trained, top-notched, basic research, physiological psychologists. *The Journal of Physiological Psychology* is one of the best journals we put out in APA, it's been losing money for years. No one has seriously suggested that we discontinue it or cut it, but it's been a losing deal, downhill, downhill. This year a sudden interest in hiring more physiological

psychologists has begun to occur. Where, in the research departments? No. In the departments of clinical psychology. Neural psychology has started to heat up. Everybody is getting excited about diagnostics using neural-psychological techniques. The most outstanding technique for neuralpsychological evaluation is the Reitan-Halstead. It's been 25 years in development, it takes years to learn it well, it takes about six and one-half hours to administer it. There's probably not a physiological psychologist one year past his PhD who couldn't sit down, with a couple of laboratory assistants, spend one year and come up with a battery far superior predictively and specifically than anything that the Halstead-Reitan can do. So we're on the threshold of suddenly finding that clinical psychologist and physiological psychologists can have a very happy, and very seminal kind of marriage that's going to do everybody a lot of good. Because of the research funding, because of the great clinical interest, is going to increase, and more physiological psychologists are going to be supported in training and they're going to be more slots in the university, etc., etc. And this is why I say as we get our act together, good things will come to the behavioral sciences. Some things are already being done.

Psychiatry has been engaged in a venture of science-fiction called the DSM Three, the Diagnostic and Statistical Manual for Mental Disorders. Those of you who have gotten preliminary copies, who have a scientific background, look at it and are appalled. Adolescent rebellion is considered a medical disorder, with medical etiology. John Conger, who is of the psychologist, Dean of the Medical School, here in Colorado was sitting calmly, looking at this and the rest of us were furious at the lack of scientific base of this nosology system. And he says, "It wouldn't be so bad, except it's incomplete." And we said, "John, how can you say that?" He said, "Well, they've left out Farrah Fawcett-Majors phobia and I see nothing in here about yogurt addiction." It is a travesty, and yet supposedly a scientific document.

Well, psychologists have been on this committee, have been working with them and we continue to get rejected. Last year we said instead of begging, instead of pleading, behavioral science has been interested in nosology and taxonomy and classification for almost 100 years and we have fairly sensible data base. Why don't we form a group within APA to develop a behavioral science taxonomy of behavior. How simple. How ridiculously simple. It's been in operation for a year. The preliminary report is very exciting. We're funding them for 5 more years, not going to rush into it. We're going to get plenty of funding for it. The names of the people on this are people you know and respect, like Joe Zubin, Parlov, Morris Lore, people who have devoted their lives in behavioral sciences to taxonomy and nosology, very, very good. I told you about the CHAMPUS studies. We have finally gotten the DOD to be concerned about the quality of the delivery of clinical services to children and adults through CHAMPUS. And there is a set of fairly massive studies on effectiveness and quality of delivery being conducted by the American Psychiatric Association; they have a grant, and by the American Psychological Association. We have hired a first rate project director, a fellow who has very successfully completed a number of behavioral science projects. The staff are all well-known scientists in psychology. Psychiatry has begun examining records and they're using the traditional patient is mildly, moderately, or severely disturbed; the patient has improved not at all, somewhat, has made major improvement; and all that other science-fiction in so far as scaling is concerned. And at the end of two years the reports will be submitted. I have no question of how much good the report from the American Psychological Association project is going to do for the behavioral sciences. From the little APA, I have no idea, I haven't seen their research design. Some of the people who have been working on DSM Three are invovled in that, which I think is unfortunate.

One of the places that we have been doing some good, as I told you, is in public information and the legislature. The APA has suddenly become more aware that the things we say to people go down on paper, particularly in Washington, and they come back to haunt you. So you got to be very careful about what you say. You got to follow the old-fashioned admonitions of stick to-your-data-base, know what your saying, and be fairly sure it's replicable. And that has done an enormous amount of good.

But the most important good for psychology in the 1980s, unquestionably has been done by the undergraduate teachers of psychology in the United States. The undergraduate departments of psychology, whatever they have done, have arranged that psychology is the single most popular elected subject in American universities. And let me tell you where that impact has had its effect.

The first time I visited a Congress person or Senator was about 15 years ago. I was introduced to Senator Strom Thurmond of Carolina. He clapped me on the back and nearly broke my clavicle. He's an enormously powerful guy in spite of his age. He says, "Son, I'm glad to see another Southerner. I sure don't talk easy with these Yankees. Now tell me how you're different from them psychiartists who are always traipsing around here wanting different things from us." And, I saw that I had my work cut out for me and I had some material. But, you know, it's and uphill battle to spend an hour with a Senator and explain what psychology is, and so on and so forth.

First time I visited Senator Inoye, which was about two and one half years ago, came into his office, sat down, gave the lefthanded shake. You know he lost his arm in the Italian Campaign. He sits down and says , "You know, I had two courses in psychology when I was at the University of Honolulu, two of the best courses I ever had. Do you know Professor so and so, do you know Doctor so and so? Wonderful teachers, really enjoyed myself. Got an idea. By the way, is Skinner still alive? Is Maslow still operating? Sit down, have a seat, we'll get some coffee." All of a sudden, this very exciting feeling from the pit of my stomach came up. For the first time, dealing with a power broker, here's a guy who didn't have to be sold anything. He took two courses in psychology as an undergraduate, he remembers the names of some of the leaders in the field, he had a good time, he's glad to see us, and he wants to know what he can do for use because he has confidence. And in the last two years this is what we've been seeing among the young Congress persons. They took psychology as undergraduates, it helped them in law school, they wish more people would take psychology, what can we do for you? So in many of the very important kids of changes, in terms of positive support for behavioral science-psychology, we owe a debt of gratitude to the undergraduate professors who influence these very important people at the time they were undergraduates and in their 20's and sometimes even earlier.

There are some things we're not doing very effectively. We're not really asking hard questions about the effectiveness of our applied work. We're more interested in finishing the project, turning in the final report, and hoping nobody finds anything wrong with it. No one should do a report for DOD without putting in a section about next steps. You learn that as a graduate student. And this research must be viewed in light of various circumstances in terms of its application, don't expect miracles. We're afraid of the qualifers that we really ought to attune our nonbehavioral science colleagues to. If you expect miracles, you're going to get con men and car salesmen doing your research. You've got to support basic research down the line. You've got to stick with us over a period of time. You have got to give us some kind of linkage between administrations and over funding periods, because we cannot guarantee results in one year, two year, three year sequence. And I think we're getting to a position of power in respect to what we've already done, where we could really make these demands.

We're doing a project on leadership with TRADOC. And I can't tell you much about it, it's still restricted. It has to do with leadership at a lower level. The said report is due by 31 July. Sure we'll give you a report, but the project will be about one-third down the road. Well we can't do that. Well you've got to, this is too important. We've had some excellent reliability checks on our instruments and we

have at least four different approaches to validity that look good. We may finally have an instrument that measures leadership at a lower level against down-the-road criteria, such as culpable incidents that require court-martial activity; wastage in storage space; defective units being turned in; real measurable kinds of good and bad leadership. You can't ask us to finish it up. We got to. So, we took a chance and we put a phone call to some of the generals that we worked with seven years ago when they were colonels. Well, you don't do that. Well you don't, except you do. It depends on who the general is, and what his attitude is about behavioral science. And, lo and behold, we're going to do some down-the-road evaluation next year on primary leadership. Never could do that before. This is a threshold, a new situation, where behavioral sciences can bring to the attention of those people who've had good experiences, that it is worthwhile to take a long range view about research and its impact.

One of the places that we haven't done the work that we could do, is in the whole new exciting area of quality of life. Last year, one of the papers, I mean two years ago, in your very own symposium, if you look in the proceedings, it was the very first paper after the keynote address, was on the quality of life. And it had to do with how you feel about being in the service. And that's a start. The quality of life research has gone well beyond that. And if you've been reading last year's literature, you'll see it appear in *Professional Psychology, American Psychologist, Quarterly Review of Research, in Evaluation.* But the area that we've skirted away from, that you could get millions in support for if you once started, and probably the guy leading the pack to give you the money would be William Proxmire, and that is the question of quality of life in productivity, not just on how you feel. But we have really avoided the question of why are we spending more in time and money and getting less: service, automobiles, refrigerators, anything you can put your finger on will cost you more and it works worse than it use to work.

I flew into Colorado Springs yesterday and the lady asked me what I wanted and about 20 minutes later I said could I have what I asked for? She said, "Did you ask for that, I forgot." And so she went and brought it and brought the wrong thing. And she took my hangup bag and I get to Colorado Springs and there's no hangup bag. And I look around, where's my hangup bag? And they're about to take off, and Colonel Sharp is thinking, "Blau didn't come." Finally, we found that she shoved it in the Denver bin instead of the Colorado Springs bin—this is the same lady. She says,"I thought you were going to Denver." I said, "Well, I said Colorado Springs, and you wrote it on the tag." And she says, "I must have put it in the wrong place."

How many of you have not had this experience. Everyday in the past year. This costs enormous money. It can be the ruination of an economy for people not to work effectively to produce. You in the military are somewhat use to this. You say, "Gee, why aren't our enlisted personnel as enthusiastic, as Gung-ho, as task oriented as they used to be?" You don't know how bad it is until you get out in the economy, or maybe you do know. And we have not addressed the question vigorously or at all for that matter of why the quality of life has decreased and decelerated because of the lack of quality of personal productivity. We talk about leadership and motivation, but we could junk both of those, if we could make advancements in finding out why some people are able to take personal responsibility for what they do at all levels and why other people are not able to and find ways of influencing this in a positive measure.

If there are answers to the question of why people aren't doing good and consistently doing good, in terms of meeting their own potential, it is likely if there are such answers in measurable, behavioral terms, it's the behavioral scientists who are going to divide the kinds of research projects to identify these factors. We are on the threshold of important times and very important things for behavioral science. The Department of Defense facility for conducting such research and application could be one of the foremost proponents of the effectiveness of what we do. You certainly know the condition of the school systems today, so learning technology is a matter of great importance. There's no major learning technology project being done in the civilian component, because the teachers absolutely don't want to hear about it. They want more money, more bricks, more mortar, but they don't want technology. But if you go down to Fort Monroe, and look at what their doing in educational technology, you realize it is the Department of Defense that is the major supporter of research in educational technology today and it's working. Some of the data is outstanding and some of you involved in it, know about it. From the Department of Defense, we could have the kind of the data that could lead to a most significant era for both application and research in behavioral sciences. But we are approaching a Franklinian conflict, because with all the opportunities and all the good work that has been done before, we have become large enough, powerful enough, and diverse enough that if our voices are clang and noise, instead of signal, we're going to spoil it not only for ourselves, but those who should receive the benefits of our work. And when I say we're in a Franklinian situation, I mean very specifically the Ben Franklin situation. Ladies and gentlemen, if we do not hang together, most assuredly we will hang separately. Thank you.

PANEL DISCUSSION:
Organizational Development in the DoD: Process or Structure?*

Organizational Development (OD) is a rapidly growing behavioral science/management approach for solving organizational problems and dealing with change. The panel examined the application of OD in the Department of Defense, and addressed the viability of process vs structural approaches.

Co-Chair: Lt Col William E. Rosenbach and Capt William H. Clover,
Department of Behavioral Sciences and Leadership, USAF Academy, CO.

Discussant: Brigadier General Joseph Kovarick, USAFR
East Stroudsburg College, PA

Members: Capt James Donovan, USN
Human Resources Management School
Millington, TN

Dr. Laurie Broedling
Navy Personnel Research and Development Center
San Diego, CA

Lt Cmdr Ray Forbes, USN
Dept Administrative Services
Navy Post Graduate School
Monterey, CA

Lt Col Frank Burns, USA
Dept of the Army/DAPE-HR
Washington, D.C.

Lt Col Dennis Umstot, USAF
Air Force Institute of Technology
Wright-Patterson AFB, OH

Captain Rick Sayre
Organizational Diagnostic Division
Human Resources Office
Fort Carson, CO

*Manuscript not available for publication.

DEVELOPMENT OF AN ORGANIZATIONAL ASSESSMENT PACKAGE

William H. Hendrix

Air Force Human Resources Laboratory

Introduction

Within organizations, management is concerned with how well their organization meets their objectives. These objectives are frequently measured in terms of productivity, cost savings, and retention of personnel. The Air Force continues to be vitally concerned with this area of organizational effectiveness. This paper focuses on the development of an organizational assessment package in support of a current Air Force program designed to improve the effectiveness of organizations throughout the Air Force.

Background

The survey instrument package under development is entitled the Organizational Assessment Package (OAP). It is being developed for use by the Air Force's Leadership and Management Development Center (LMDC) Maxwell AFB, Alabama. LMDC's objectives include: (a) providing consultative service to Air Force commanders, (b) providing Leadership and Management training to Air Force personnel in their work environment, and (c) performing research in support of (a) and (b) above. The consultative role involves organizational problem area identification and recommendations for reducing or eliminaing problems identified. The OAP is being designed to meet these objectives of the Leadership and Management Development Center. First, the OAP will provide a means of identifying existing strengths and weaknesses within organizational work groups, and aggregated work groups, such as directorates. Second, research results can be fed back into their Professional Military Education, other leadership and management traning courses, and, when action is required, to Air Staff and functional Offices of Primary Responsibility (OPRs). Lastly, the OAP data base established can be used for research to stregthen the overall Air Force organizational effectiveness program.

Three Component Organizational Effectiveness Model

The Three Component Organizational Effectiveness Model[1] (c.f. Figure 1) previously reported by Hendrix (1976) considered Organizational Effectiveness (E) to be a function of: the criterion selected (c); the managerial style employed (m); and the situational environment (s), which includes the manager's subordinates, peers, and other personnel in the environment. That is: $E = f(c, m, s)$.

The Organizational Assessment Package (OAP) was designed to measure the basic components of the Three Component Organizational Effectiveness Model. As can be noted in Figure 1, the Supervisory Job Inventory (SJI) was designed to measure managerial style (m), while the situational environment (s) was to be measured by two sections of the OAP, the Background Information section and the Organizational Job Inventory (OJI). The criteria selected included satisfaction, organizational climate, and perceived productivity. These were to be measured by the sections entitled: Job Satisfaction Questionnaire (JSQ), Organizational Climate Inventory (OCI), and Perceived Productivity Index (PPI). Hard data when available will be collected separately and merged with the OAP data base in future studies.

OAP Factors

Items within each of the OAP sections have been written to measure certain factors. The *Background Information* section contains biographical information items and items associated with factors in the situational environment. The factors in the situational environment which the items attempt to measure include: (a) organizational level of work group, (b) work group type, (c) work group size, (d) group member maturity, (e) organization's geographic region, (f) extent to which work group meetings are used to establish goals, (g) extent of communication between work group members, and (h) stability of work hours. In addition, the situational environment is in part measured by the Organizational Job Inventory (OJI). In the main, the factors included in the OJI are based on the job enrichment model proposed by Hackman, Oldham, Janson, and Purdy (1975). They proposed five basic factors which they called Core Job Dimensions. Those were: (a) skill variety, (b) task identity, (c) task significance, (d) autonomy, and (e) feedback from the job. These factors are to be measured by the OFI plus one additional work related factor which is labeled *Work Interference*. This factor deals with the extent and adequacy of: (a) additional duties, (b) equipment and supplies, and (c) provided work space.

In the criterion area, organizational climate is measured by the *Organizational Climate Inventory* which includes the factors of: (a) communications, (b) general organizational conditions, (c) employee concern, (d) employee commitment, (e) decision making, and (f) recognition.

Another criterion area is that of job satisfaction which is measured by the *Job Satisfaction Questionnaire* (JSQ). This questionnaire contains 30 items which are descriptions of 30 factors out of 35 factors isolated by Gould (1975) in an unpublished study. The methodology and items used to isolate the factors can be found in Tuttle, Gould, and Hazel (1975). The 30 factors are listed in Table 1.

The last criterion is perceived productivity and is measured by 4 items contained wihtin the *Perceived Productivity Index* section. The items measure perceived productivity in terms of the work group's: (a) quantity of work output, (b) quality of work output, (c) performance when high priority work arises, and (d) whether flow of work to or from the work group is impaired.

The *Supervisory Job Inventory* (SJI) consists of 81 items relating to supervisory behavior. Once an adequate sample has been obtained, these items will be factor analyzed and the resulting factors will be used to depict differing managerial behaviors.

Method

A small scale study ($n = 144$) was conducted at Lackland Air Force Base during May 1977. One purpose of the study was to collect critique information of the OAP in order to improve it. In addition, the data served to provide an initial baseline in terms

of means and standard deviation for each item on the OAP. An intercorrelation matrix consisting of the OAP item variables plus a series of variables generated from the original variables, was used to: (a) delete items which did not intercorrelate well with the stated factors, and (b) establish simple correlational relationships between variables in the situational environment and managerial area with criteria items. The OAP previously described is the result of revisions based on data collected from the Lackland study.

In February 1978 a second study (n = 174) was conducted at a Major Air Command headquarters using a revised version of the OAP. Items common to both OAP versions are discussed in this paper.

Results

The major modification due to the Lackland study was the deletion of the Job Diagnostic Survey (JDS) (Hackman, et al, 1975) from the instrument package, with the OJI being used instead to establish the job enrichment variable values. The reason for deleting the JDS instead of OJI was to reduce the total pages in the OAP (i.e., the JDS is approximately 7 pages and the OJI is 2 pages) and to have the format of the instrument the same as that of the other instruments within the OAP. The JDS is an excellent instrument and if the OJI indicates a job enrichment problem exists within an organization, then a more thorough examination could be accomplished using the JDS. Table 2 lists the intercorrelations between the job enrichment factors on the JDS with their counterpart on the OJI. Table 3 presents for the Lackland study the intercorrelations of selected criterion items with the situational variables of: (a) a total score across items on the OJI (OJI Total), (b) the Motivation Potential Score (MPS) and Growth Need Score (GNS) as defined by Hackman et al, (1975), (c) the Need for Enrichment Index (NEI) which is derived from the OJI and is the total score of all items indicating a need for enrichment, and (d) the Job Motivation Index (JMI) which is computed with the same formula as the GNS. Table 3 also presents for the Major Command study (MAJCOM study) the intercorrelations of criterion items with the situational variables of: (a) OJI Total, (b) NEI, (c) JMI, and (d) JMI-Additive. The JMI is a multiplicative model computed by the formula:

$$\frac{\text{(skill variety + task identity + task significance + work irritants)}}{4}$$

$$\text{X Autonomy X Job Feedback.}$$

The JMI-Additive index is computed by the same formula except that autonomy and feedback are added instead of being multiplied.

Discussion

The OAP is designed to provide indicators of the manager's behavior, the situational environment, and criteria of organizational effectiveness. Should problem areas be identified, then a more detailed investigation will be performed by on site consultation teams. Once validated, the OAP should provide a means for: (a) identifying organizational strengths and weaknesses, (b) establishing appropriate managerial behavior in different situations with different criteria of success, and (c) identifying and resolving functional, career field, or systematic Air Force problems.

The Organizational Assessment Package (OAP) is presently being validated based on data collected at Air Force installations. The proposed validation phase extends from January 1978 to July 1978. Based on research results, the OAP will be revised and turned over to LMDC for operational use by their travel teams. Data obtained from the operational use of the OAP will be analyzed to further improve the effectiveness of LMDC's consulting program.

FOOTNOTES

[1]In Hendrix (1976) the model was initially entitled the Three Component Leadership Effectiveness Model and has since been expanded to focus on the entire organization.

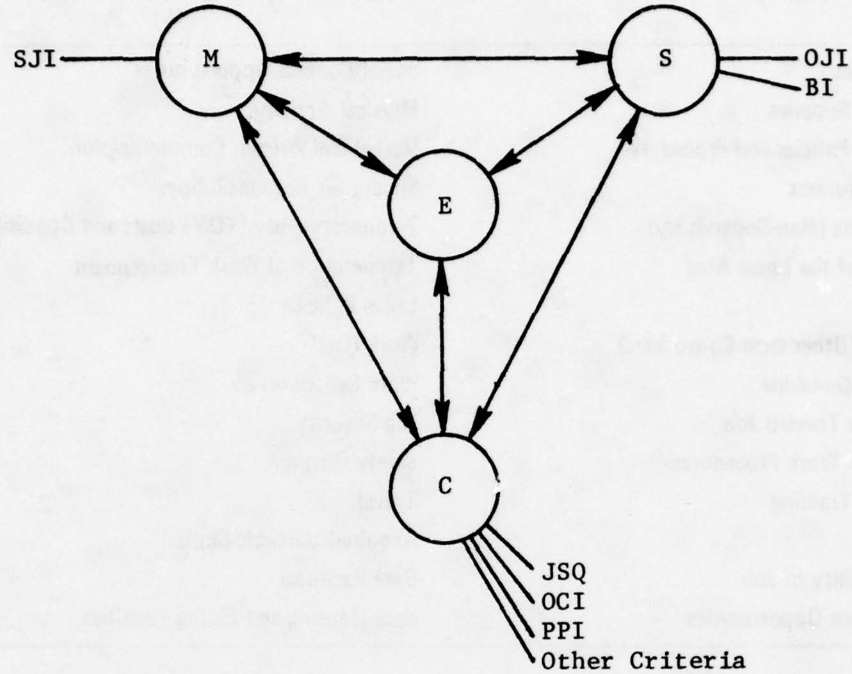
REFERENCES

Gould, R. B. *Dimensions of Job Satisfaction: Development of the Air Force Occupational Attitude Inventory.* Unpublished AFHRL research project, 1975.

Hackman, J. R., Oldham, G., Janson, R., & Purdy, K. A new strategy for job enrichment. *California Management Review,* 1975, *17* (4), 57-71.

Hendrix, W. H. *Contingency Approaches to Leadership: Review and Synthesis.* AFHRL-TR-76-17, AD-A028 485. Lackland Air Force Base, Texas: Occupational and Manpower Research Division, Air Force Human Resources Laboratory, June 1976.

Tuttle, T. C., Gould, R. B., & Hazel, J. T. *Dimensions of Job Satisfaction: Initial Development of the Air Force Occupational Attitude Inventory.* AFHRL-TR-75-1, AD-A014 796. Lackland AFB, TX: Occupational and Manpower Research Division, Air Force Human Resources Laboratory, June 1975.



Legend:

M = Management style
S = Situational environment
C = Criterion
E = Effectiveness
SJI = Supervisory Job Inventory
JDS = Job Diagnostic Survey
OJI = Organizational Job Inventory
BI = Background Information form
JSQ = Job Satisfaction Questionnaire
OCI = Organizational Climate Inventory
PPI = Perceived Productivity Index

FIGURE 1. THREE COMPONENT ORGANIZATIONAL EFFECTIVENESS MODEL

TABLE 1

JOB SATISFACTION QUESTIONNAIRE FACTOR ITEMS

Additional Duties	Social Contact Opportunities
Equipment and Supplies	Physical Activity
Information on Policies and Procedures	Verbal and Written Communication
Feeling of Helpfulness	Supervisor Responsibilities
Control of Others (Non-Supervisory)	Temporary Duty (TDY) Costs and Conditions
Characteristics of the Local Area	Temperature of Work Environment
Work Space	Leave Policies
Social Contact (Other than Co-workers)	Work Itself
Co-Worker Relationships	Work Schedule
Family Attitude Toward Job	Job Security
Independence in Work Procedures	Safety Program
Job-Associated Training	Travel
Job Hazards	Acquired Valuable Skills
Moral Acceptability of Job	Base Facilities
Self-Improvement Opportunities	Base Housing and Eating Facilities

TABLE 2

CORRELATIONS OF JOB ENRICHMENT FACTORS ON THE JDS WITH THOSE ON THE OJI

FACTOR	CORRELATION COEFFICIENT
Skill Variety	.60
Task Identity	.60
Task Significance	.64
Autonomy	.44
Job Feedback	.59

TABLE 3

SITUATIONAL ENVIRONMENT VARIABLE PREDICTORS OF CRITERIA

	CRITERIA						
	Work Satisfaction	Perceived Productivity[a]				Climate[b]	
Situational Variable		1	2	3	4	1	2
LAFB Study							
OJI Total	.52	.26	.41	.30	-.12	.26	.22
MPS	.62	.21	.39	.32	-.21	.26	.22
GNS	.16	.21	.15	.16	.10	.14	.22
NEI	.23	.16	.19	.16	-.04	.18	.16
JMI	.56	.32	.44	.32	-.17	.30	.25
MAJCOM Study							
OJI Total	.56	.22	.28	.31	-.19	.38	.34
NEI	.19	.11	.22	.26	.06	.05	.26
JMI	.47	.15	27	.29	-.21	.34	.31
JMI-Additive	.52	.18	29	.32	-.23	.41	.35

[a]Perceived Productivitiy

1 = Quantity of Work Output
2 = Quality of Work Output
3 = Performance when high priority work arises
4 = Efficiency in work flow from and to work group.

[b]Climate

1 = You are proud of organization
2 = You feel responsible for your organization

RESEARCH DESIGN FOR THE EVALUATION OF THE AIR FORCE LOGISTICS COMMAND-WIDE ORTHODOX JOB ENRICHMENT (OJE) PROGRAM

Russell F. Lloyd

HQ Air Force Logistics Command

Abstract

In September of 1976, the Commander of Air Force Logistics Command made a decision to implement the behavioral science intervention of Orthodox Job Enrichment (OJE) across his Command. In so doing the Headquarters, five major depot-level maintenance facilities (each employing approximately 20,000 employees), the Air Force Acquisition Logistics Division, two minor maintenance facilities and a hospital all became subject to participation. In short, approximately 92,000 personnel across the nation became potentially involved. The OJE Program organization involving nearly 150 personnel located at varying geographical sites across the nation implement administer the program on a day-to-day basis.

Separate from the OJE Program, however, is a *built-in* evaluative capacity. This evaluation is an "independent" one conducted by a HQ AFLC organization (HQ AFLC, DCS Plans and Programs (XR)) which is apart from, or outside of formal OJE program involvement. The implicit purpose of this organizing theme is to maintain a detached, objective point of view.

The OJE Program within AFLC is one of the largest organizational change interventions ever attempted and, concomitantly, the largest ever to be evaluated in rigorous professional fashion.

INTRODUCTION

In September of 1976, the Commander of the Air Force Logistics Command (AFLC), United States Air Force, made a decision to implement the behavorial science intervention of Orthodox Job Enrichment* across his Command. In so doing, the Headquarters, five major depot-level maintenance facilities (each employing approximately 18,000 employees), the Air Force Acquisition Logistics Division, and two minor maintenance facilities all became subject to participation. In short, approximately 92,000 personnel across the nation became potentially involved. The OJE Program organization involving nearly 150 personnel located at varying geographical sites across the nation implement administer the program on a day-to-day basis.

The AFLC Program Plan 76-23, September 1976 (the basic directive guiding the AFLC OJE effort), clearly provides for ". . . two separate modes of measurement and evaluation". The first is what has been called an "internal evaluation" which is to be conducted by those personnel who are a part of the OJE Program; specifically, the administrators and keymen. *Separate* from this internal evaluation is a built-in evaluative capacity. This second mode of measurement and evaluation, however, is an "independent" one conducted by a HQ AFCL organization which is apart from, or outside of formal OJE program involvement. The implicit purpose of this organizing theme is to maintain a detached, objective point of view. The Deputy Chief of Staff Plans and Programs (XR) has been designated as *that independent, objective organization.*

The AFLC Command-Wide OJE Program is perhaps the largest organizational change intervention ever attempted in the history of the field. Certainly, it is the largest job redesign intervention ever attempted. Concomitantly it is the largest intervention ever to be evaluated *in a scientifically rigorous fashion.*

SCOPE OF AFLC COMMAND-WIDE OJE PROGRAM

The Command-Wide OJE Program is monitored at HQ AFLC by the Deputy Chief of Staff, Personnel. It is administered on a day-to-day basis by a HQ detachment located at Hill AFB, Ogden, Utah. That location provides easy access to Herzberg and Associates and to Hill personnel who have considerable experience in the OJE Program by virtue of their involvement as the pilot project.

The detachment contains 7 personnel who supervise the daily administration of the OJE Program at the Headquarters, the Air Force Acquisition Logistics Division, the 2750th ABW, AGMC, MASDC,and all five Air Logistics Centers. At each field location, there is, in turn, an OJE Program Office which consists of approximately 4 full-time personnel who administer the OJE Program at that location. Working with them are various numbers of keymen (avg: 17) in each directorate. These individuals are the ones who actually conduct the OJE intervention at the project level.

At this writing there are approximately 150 full-time personnel involved in the OJE Program across the Command. To date, 176 projects have been initiated are in process are completed affecting in excess of 10,000 employees. A sum of three million dollars is expected to be invested on an annual basis.

DATA COLLECTION INSTRUMENTS

Data is being collected via five techniques. First, an AFLC Quality of Worklife Survey was developed and approved at HQ AFLC, HQ USAF and AFMPC. It consists of Hackman and Oldman's (1974) Job Diagnostic Survey, Mott's (1972) Self-Perception of Organizational Effectiveness (SPOE) index, and two indices developed by the author to measure leadership style and communication pattern (quantity and quality). Second, similar quality of worklife and productivity data will be collected via the utilization of the Critical Incident Interview format used by Herzberg as an exclusive data collection device. A comparative analysis of findings between the above two will be conducted. Third, a customer survey will be conducted wherein perceptions of the "customer" will be collected with regard to changes in quality

---

*Orthodox Job Enrichment (OJE) is a registered trademark of Herzberg & Associates.

and quantity of services rendered. Fourth, an effort is underway to devise, where possible, objective (hard) productivity measures for each OJE project control group evaluated. Finally, a historical log will be maintained jointly by the supervisor and the field evaluator wherein information is recorded on any event which in some way could impact on the effect of the OJE intervention.

RESEARCH DESIGN

Field researchers, it is claimed, unnecessarily acquiesce to the characteristics of the organization, or to the pressures of management, or to other realities of the field setting when designing their research paradigms. Such a blanket criticism is a generalization, to be sure, but to the extent it is true this author wanted to utilize as pure a design as was feasible given the actual limitations field research *does* present.

Using research notation suggested by Campbell and Stanley (1963), Figure 1* presents a diagram of the research design. The features of the design are threefold: random selection and assignment (partial) of work groups, the utilization of before-after (pre-post) measures, and the utilization of control groups.

This paradigm very conveniently generates data to control for all of the sources of internal invalidity as discussed by Campbell and Stanley (1963). Randomization itself controls for regression and selection effects. Maturation effects are controlled for in that they should be manifested equally in experimental and control groups, and effects due to history are controlled for insofar as general historical events that might have produced an 0_{1_A} -- 0_{2_A} -- 0_{3_A} -- 0_{4_A} difference would also produce an 0_{1_B} -- 0_{2_B} -- 0_{3_B} -- 0_{4_B} difference. The effects of testing are controlled for by the way of the comparison 0_{1_B} -- 0_{2_B} with 0_{4_C}. Instrumentation is controlled for where the 0's are achieved by responses to a fixed instrument. Data are also available to discern whether mortality offers a plausible explanation of any 0_{1_A} -- 0_{2_A} -- 0_{3_A} -- 0_{4_A} gain. Comparisons of the quantity of respondents at 0_{1_A} versus 0_{2_A}, 0_{3_A}, 0_{4_A}, or 0_{1_B} versus 0_{2_B}, 0_{3_B}, 0_{4_B}, will provide at least a surface indication mortality. Granted, differential dropout can still occur while quantity of respondents remain stable, but randomization gives some assurances that when and if mortality is ocurring the character of the dropouts is similar in both the experimental and control groups. Finally, the effects due to interaction of selection and maturation, etc., are controlled for via randomization (selection) and with the comparison of experimental groups to control groups.

CONCLUSION

The objectives of the OJE Evaluation Program are threefold: *to gather* a representative sample of data pertinent to the OJE Program objectives (*i.e., improve quality of worklife and enhance productivity*), *to evaluate* that data, and *to provide information* to the Command to utilize in evaluating the long-term utility of Orthodox Job Enrichment. In so doing, HQ AFLC can be provided conclusive information with regard to the effects of the Command-Wide OJE Program. Data-based decisions can be made concerning strengths of the OJE Program that should be continued and areas needing further evaluation. Contributions can be made *beyond the area of AFLC. The Air Staff, the* DOD, private industry and, then too, the professional community can *all* learn valuable insights from the results of this rigorous evaluation of a concomitantly large-scale OJE program. The Air Force Logistics Command has indeed the potential to be a "precedent setter".

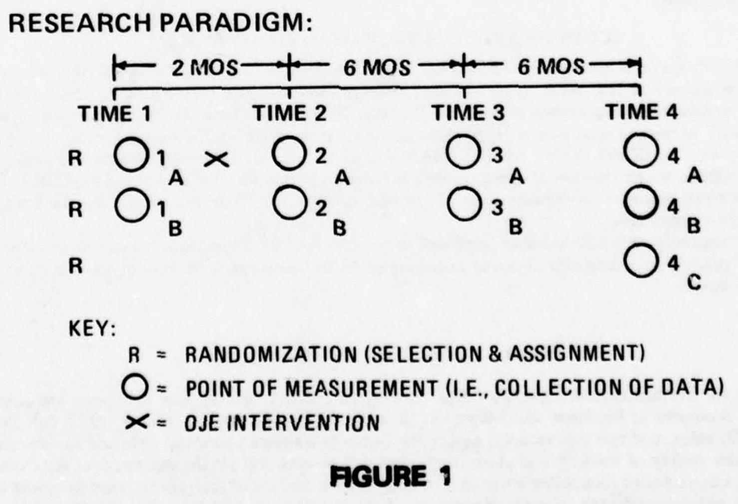


FIGURE 1

*R represents the random selection and assignment of work groups, O represents some process of observation or measurement (i.e., the pre-and Post-tests), and X represents the exposure of a work group to an experimental variable. The X's and O's in a given row are applied to the same work group, the left-to-right dimension indicated temporal order, and the X's and O's vertical to one another are occurring simultaneously.

BIBLIOGRAPHY

Campell, D. T., and Stanley, J. C. *Experimental and Quasiexperimental Designs for Research.* Chicago: Rand McNally, 1963.

Hackman, J. R. and Oldham, G. R. The Job Diagnostic Survey: An Instrument for the Diagnosis of Jobs and the Evaluation of Job Redesign Projects. Technical Report No. 4, Department of Administrative Sciences, Yale University, 1974.

Mott, P. E. *The Characteristics of Effective Organizations.* New York: Harper and Row, 1972.

AN EVALUATION OF ORGANIZATION EFFECTIVENESS: AN INVESTIGATION OF THE EFFECTS OF SURVEY FEEDBACK AS AN ACTION RESEARCH INTERVENTION IN THE U. S. ARMY

Jerome Adams

United States Military Academy

ABSTRACT

This study reports the results of a six month action research project designed to evaluate the effects of survey feedback used as an intervention strategy within engineer units in a military setting.

The primary dependent variables included: (a) unit efficiency, (b) measures of soldier attitudes, both general and specific satisfactions, (c) intergroup relations, (d) supervisory consideration.

It was hypothesized that the experimental treatment units would have a significant improvement in organization effectiveness following the survey feedback intervention. It was also hypothesized that soldiers within the experimental units would have improvements in their levels of work satisfaction. Additionally, it was hypothesized that the survey feedback units would have better intergroup work relations. Finally, it was hypothesized that soldiers within the experimental units would perceive greater supervisory consideration for soldiers' work related problems.

Conclusions about evaluating organization effectiveness programs in work settings were also discussed. Finally, recommendations have been made about the practical implications of this study in the U.S. Army.

The purpose of evaluation research is to measure the effects of a program or set of related activities against the goals the program set out to accomplish as a means of contributing to decision making about improving future programming (Weiss, 1972). The need for evaluation of organization development interventions and programs is often mentioned in the literature (Argyris, 1970; Armenakis, 1973; and Kahn, 1974). However, with the exception of the work of Meyers (1975), a review of the leading journals, reveals that there is a paucity of reports of evaluation studies of well defined programs, directed toward clear decision making functions.

This paper reports the results about a classical evaluation research effort of an organization development program by reporting the effects of a data based intervention intended to improve unit effectiveness in the United States Army.

RATIONALE OF STUDY

During the period 1973 through 1975, the United States Army conducted a pilot program designed to adopt proven management concepts and practices, which when incorporated within military units, would improve the effectiveness of these units. This program included such organization development activities as job enrichment, participative management, team building, and survey feedback. One tangible result of the pilot effort was the creation of the Army's Organizational Effectiveness Program.

Three goals of the organization effectiveness program included:

- — to increase individual and unit effectiveness
- — to create an organizational commitment to which soldiers are actively involved in planned actions to improve the unit's performance in meeting its mission of being combat ready
- — to strengthen the chain of command

Organization effectiveness* activities have been conducted within military units by Army trained internal consultants. This study was part of a larger still ongoing evaluation research project commissioned to compile empirically based data regarding the effectiveness of organization development programs and activities in military settings. The focus of this study was to evaluate the effects of a survey feedback program against the goals it set out to accomplish. Specifically, did the survey feedback program produce outcomes of improved work efficiency, job satisfaction and better working relationships—and in turn did these changes improve individual and unit effectiveness?

DESIGN

The design of the study involved four units from within the engineer organization. All units were given a pre-test measure at the time. Two of the four units were randomly selected to participate in the survey feedback program; one unit was designated as a placebo control, and the fourth unit was designated as a no treatment control.

FINDINGS

The four dependent variables were selected in this evaluation research project.

- — Work Project Efficiency
- — Affective Responses to Work
- — Intergroup Relations
- — Supervisory Consideration

In terms of work product efficiency, statistically significant improvements did occur in the two units which participated in the survey feedback programs for one of the three types of projects measured. Improvement was also noted in the placebo unit for the same type of project. This positive outcome may be a result of better work relationships which were the result of race relations seminars. No change occurred in the control unit's level of efficiency. There was no efficiency improvement in any of the engineer units for type two and type three projects.

*The term Organization Effectiveness is used in the Army to denote Organization Development.

Regarding soldiers' affective responses to work, two statistically significant findings were noted. First, in the criterion measure of General Satisfaction, there was a statistically significant decline for one of the two units which participated in the survey feedback intervention. The measured decline occurred soon after that unit commander unexpectedly announced that he was resigning from the Army. Improvement was noted in the second unit participating in the survey feedback intervention, but the measured increase was not statistically greater than the improvement noted in the placebo unit. The tool unit had a negative change in General Satisfaction. The second positive change noted in soldier's attitudes was found in the Specific Satisfaction measure. Both of the survey feedback unit improvement was statistically greater than the control conditions.

No statistically significant changes occurred among the units when Intergroup Relations was measured. One survey feedback unit improved; the other declined but not significantly so. Some improvement in Intergroup Relations was also noted in the placebo unit but this was probably due to the better communications facilitated by the race relations meetings. No improvement took place in the control unit.

Statistically significant changes occurred in the two survey feedback units regarding one of the two measures of Supervisory Consideration. One unit participating in the survey feedback program experied a significant decline while the other survey feedback unit's scores improved significantly for the variable—Consideration Given (NOW). The fact that significant changes occurred only within the two survey feedback program units is testimony that meetings were being held. However, the contrast in outcomes of the meetings appears to suggest that both groups were not equally satisfied with the amount of consideration which was given. Thus, we conclude that one unit participating in the survey feedback program chose to deal with the issue of a lack of perceived supervisory consideration and as a result, this unit showed an improvement on the post-test. However, the leadership style of the second survey feedback program leaders caused soldiers to perceive these leaders as less supportive. That is, this unit's leaders held feedback meetings but limited the involvement of the soldier's in the discussion and planning of how to deal with problems raised. Thus, these soldiers reported a lack of Supervisory Consideration Given (NOW).

In the second measure of Supervisory Consideration (Wanted), there were no statistically significant changes among the units over time.

ALTERNATIVE CONDITIONS NECESSARY TO OPTIMIZE DESIRED OUTCOMES

The results reported in this study followed a classical evaluation research approach. That is, the intent of the research was to measure the effects of survey feedback against the goals it set out to accomplish (i.e., to improve work efficiency, job satisfaction, and better work relationships). The study reported the results as they actually occurred.

Now, learning from this study, there are four changes needed to demonstrate the positive effects of survey feedback programs in military settings. These four changes serve as practical recommendations to enhance the positive outcomes that survey feedback can provide.

The first recommendation is to isolate units which will be involved in the research. Archer (1975) describes this isolation approach as a "sheltered experiment." That is, the object is to throw a shelter around the organization, which will allow the survey feedback process to develop in the manner called for in the waterfall design. In contrast to the study just reported, unintended outside requirements, additional work demands, and atypical work support duties would not be imposed upon units involved in the survey feedback program. The isolation would significantly enhance the probability of diagnosing meaningful work problems, which the participants themselves can change.

Isolation would also include the elimination of outside surveys and testing. That is, the units involved in survey feedback would not be subjected to other Army surveys during the survey feedback schedule. One of the important phases of survey feedback is the opportunity for the participants to own the data. If the soldiers are asked their opinion in surveys which provide no intended feedback, then two possible negative effects can occur: (1) the soldiers have raised expectations about obtaining results from the survey which are not fulfilled, and (2) the soldiers do not develop trust that any improvements will result from their participation in surveys.

Finally, isolation should include stabilizing the individuals who occupy the key leadership positions in the units involved in the survey feedback process. That is, when meetings are run by the relevant work groups, stability of key persons is necessary for meaningful interaction to occur between peers and their supervisors and supervisors and their subordinates. Personnel turbulence disrupts this critical pattern of interaction and could reduce the intended success experiences which the meetings are designed to provide.

The second recommendation made to optimize the positive effects of survey feedback is to revise the General Organizational Survey used by the Army. That is, the instrument should be designed such that a few criteria can be measured effectively as opposed to trying to measure too many variables. Instrument validation is a time involved process which admitedly could take years to accomplish. Nevertheless, two general guidelines should be used: (1) the instrument should provide an opportunity for the leaders to learn what they want to know about their units, and (2) the instrument should allow the soldiers to express what they think the perceived problems of the unit are. Once these objectives are met, the survey instrument will be tailored to each organization's problems.

The third recommendation suggested here is to have two interventionists—an evaluator and an internal consultant. The evaluator would shadow the activities of the internal consultant. This means that the evaluator would observe the major intervention activities and meet with the internal consultant after each appropriate phase to discuss what had taken place, and where the unit is going in the next steps. The evaluator as a shadow offers two strengths to the intervention program: (1) The evaluator is an objective source of information, who as a disinterested person, can note some cues and patterns of behavior which the client unit members and the internal consultant may not readily see, and (2) The evaluator is an additional resource person to the internal consultant who is available throughout the intervention program.

The last recommendation to optimize the use of survey feedback in the Army is to differentiate between the intervention entry level and the intervention working level. The entry level of intervention within the Army is the battalion. Thus, the battalion commander must approve activities which are planned within subordinate elements of that organization. However, in survey feedback programs, the bulk of the data generated for action comes from the working level of the platoons and squads. Therefore, internal consultants need to spend time with first line supervisors (i.e., platoon and squad leaders), to legitimize this process to them, since these leaders are the ones who will be most intimately involved in the intervention program. Too often officers, who are the internal consultants, tend to devote their time and emphasis with the higher ranking persons within the organization. It is more desirable to devote time with top levels of authority for the entry phase, but the consultant's focus must shift to the working levels, platoon and squad leaders, to enhance the survey feedback program's chances of success. Consideration of the four alternative recommendations may aid future efforts to optimize desired outcomes of survey feedback programs in military settings.

REFERENCES

Alderfer, C.P. and Ferriss, R. "Understanding the Impact of Survey Feedback," in W.W. Burke and H.A. Homstein (Eds.), *The Social Technology of Organization Development*. Fairfax, Virginia: NTL Learning Resources Corporation, 1972.

Armenakis, A.A. Practices and Problems of OD Practitioners in Conducting Evaluations of OD Efforts. Unpublished Paper, Auburn University, June, 1973.

Archer, J.T. "Achieving Joint Organizational, Technical and Personal Needs: The Case of the Sheltered Experiment of the Aluminum Casting Team," in L. Davis and A. Cherns (Eds.), *The Quality of Working Life*. Vol 2, New York: The Free Press, 1975.

Argyris, C. *Intervention: Theory and Method*. Reading, Mass: Addison-Wesley, 1970.

Campbell, D.T. and Stanley, J.C. *Experimental and Quasi-experimental Designs for Research*. Chicago: Rand McNally, 1966.

Clark, A.W. "Sanction: A Critical Element in Action Research," *Journal of Applied Behavioral Science*. Vol 8, 1972.

Cook, T.D. and Campbell, D.T. "Experiments in Field Settings," in M.S. Dunnette (Ed.), *Handbook of Industrial and Organizational Psychology*. Chicago: Rand McNally, 1976.

Crowne, D.P. and Marlowe, D. *The Approval Motive: Studies in Evaluative Dependence*. New York: Wiley, 1964.

Fleishman, E.A. *Manual for the Leadership Opinion Questionnaire*. Chicago: Science Research Associates, 1960.

Hackman, J.R. and Oldham, G.R. *The Job Diagnostic Survey: An Instrument for Diagnosis of Jobs and the Evaluation of Job Redesign Projects*. Technical Report #4, Department of Administrative Sciences, Yale University, May, 1974.

Kahn, R.L. "OD: Some Problems and Proposals," *Journal of Applied Behavioral Science*. Vol 10, 1974.

Meyers, W.R. "The Politics of Evaluation Research: The Peace Corps," *Journal of Applied Behavioral Science*. Vol 11, 1975.

Rosenthal, R. and Rosnow, R.L. (Eds.) Artifact in Behavioral Research. New York: Academic Press, 1969.

Weiss, C.H. Evaluation Research: Methods of Assessing Program Effectiveness. Englewood Cliffs, New Jersey: Prentice-Hall, 1972.

TOP DOWN OR BOTTOMS UP IN SURVEY FEEDBACK

Russell F. Lloyd, Cpt.
U. S. Air Force
Wright-Patterson AFB

Paul D. Tolchinsky
Krannert Graduate School
Purdue University

The focus of this paper is on the human processual intervention technique of survey feedback. Although no authoritative volume has yet been written about this developmental tool (Bowers & Franklin, 1972), survey feedback techniques are said to "value human fulfillment highly and expect improved organizational performance to follow on improved human functioning and processes" (Friedlander & Brown, 1974, p. 325). The technique is based upon the rationale that data which indicated a discrepancy between the present state of affairs in the organization and the desired state are motivators of change, much as the process of differential perception provides motivation for change within the individual (Peak, 1955).

Over the past few years different approaches to feedback in organizations have been developed. Each approach purports to be a systematic, preplanned, and tested way of structuring feedback activities. Each attempts, to varying degrees, to deal with the question of how to ensure that the process and content of feedback will be effective and that ideas will be turned into concrete action. The major approaches, or feedback designs are presented in Table 1.

TABLE 1

Major Systematic Approaches to Feedback and Action Planning

Feedback Approach	Feedback Aimed at	Assurance of Feedback to all	Consultant Present	Planning Focus	Follow-up Structure	Comments
Survey Guided Develop-ment	Family Groups	None	Yes	Top Down	Quasi	only to lower after at higher echelons: may be stopped at any time
Bottom up Develop-ment	Family Groups	Yes	Yes	Bottom up (Intra-group)	Formal	only to higher after at lower echelons: formal reports submitted to superior
Peer Group Intergroup	Peer	None	Yes	Peer/ Inter	None	vertical hierarchy in-volved after peer group meets. No specific focus
Data Handback	Manage-ment	None	No	None	None	decision to proceed made after data viewed
Collateral Group Feedback	Task Force	None	Yes	Organiza-tion	None	task group responsible for data collection, interpretation and action

Several studies purport that survey interventions can result in improved participant attitudes. Baumgartel (1959) found evidence of improved attitudes for groups at Detroit Edison. Similarly, Mann (1961) demonstrated improvements in attitudes toward work, supervision, progrsss, and ability to get the job done. Others (Bowers, 1973; Brown, 1972; and Miles, et.al., 1969) have found similar results associated with survey feedback techniques.

There is, however, little research to support the notion that improved organizational productivity is an outcome of such interventions. Hand, Estaffen, and Sims (1975), in a laboratory study involving college students and a business simulation, reported that groups receiving a survey intervention were more satisfied, but did not outperform control groups. Similarly, Kimberly and Nielsen (1975), reported improvements in product quality and employee attitudes, but no change in the level of productivity after an intervention. Finally, in his study of three intervention techniques (socio-technical systems, job redesign, and survey feedback), Pasmore (1976) found significant productivity changes with only a combined socio-technical survey feedback intervention. No gains were associated with survey feedback alone or with job design survey feedback.

Thus, the extant literature tends to support the belief that survey interventions are primarily associated with improvements in employee attitudes. This literature does not, however, (with the possible exception of Pasmore, 1976), support the notion that improving employee attitudes will precede changes in behavior. If behavior changes do occur, there is little indication that these materialize into improved organizational performance.

This research specifically compares the differential effectiveness of three of the variations (survey guided development, bottom up developent, and data handback). No treatment and post-test only control conditions are also incorporated.

In summary, to explicate the causal process, the present research tested the effects of individual and situational variables which mediate the outcomes of survey feedback. Thus far, little is known about changes in individual variables (such as expectations) that result from organizational development interventions. The longitudinal nature of this study allows such issues to be examined. In a similar fashion the effects of internal structural and functional variables (i.e., decision styles) were observed and reported.

The Experimental Setting

This field experiment was conducted at a Department of Defense military installation in the midwest. The facility employes in excess of 2,700 civilians and approximately 150 military personnel with an annual payroll of $45 million. The installation covers 56 acres and it facilities and equipment are valued at $158 million. The decision to utilize randomization represented the most difficult design question of the entire research. Survey feedback interventions have almost exclusively involved entire vertical segments of an organization, if not the entire organization itself. In so doing, randomization is simply not an issue since there are rarely enough vertical segments available in an organization for random selection. To randomize, then, one cannot involve whole vertical segments of an organization. Certainly, randomization has distinct advantages: without it one's ability to generalize is foresaken.

Intervening in the Organization

Prior to the administration of the survey pretest all work groups were randomly assigned to treatment conditions. Conditions one (traditional "waterfall"), two (data handback), three (bottoms up), and four (pretest-posttest) were surveyed initially. The activities that preceded this were the same for all units in these conditions. Once the survey had been completed treatments one, two and three received training. All group supervisors received one day's training on how to interpret the data, conduct meetings and plan actions. This program was identical for all groups. Data handback groups received training on interpreting results, additional planning and encouragement to meet with employees to feedback the data. These supervisors did not role play feeding back data.

There are several differences between these treatments that must be reiterated. Bottoms up groups were given specific forms for action planning, including kinds and types of issues that must be addressed. Additionally, action planning progressed in a sequential fashion up through the hierarchy. Traditional groups did not have specific requirements regarding issues to be addressed. This group was required to meet and submit reports to top management. Only the bottoms up groups submitted specific plans directly to their supervisors. Data handback groups were not required to meet with their employees nor submit action plans.

At this point, the researchers provided consultative skills to supervisors. In either method, during the feedback session, the consultant has the responsibility of insuring that the issues generated, by the discussion of the data, are transformed into specific proposals for change. The consultant rarely suggests what changes are needed; rather it is said that this individual acts as a "catalyst," or "transducer" during the feedback meetings (Pasmore, 1976; Bowers & Franklin, 1972, 1977).

Control groups were not involved in any part of the data feedback or subsequent training. Upon completion of this research sequence, control groups were trained in "bottoms up" feedback techniques. This feedback action planning was conducted in the exact same manner as previously described.

Thus the research design was as follows:

$$\begin{array}{ccccc} R & 0_1 & X_1 & 0_2 & 0_3 \\ R & 0_1 & X_2 & 0_2 & 0_3 \\ R & 0_1 & X_3 & 0_2 & 0_3 \\ R & 0_1 & & 0_2 & 0_3 \\ R & & & & 0_3 \end{array}$$

Where X_1 = Model II

X_2 = Data handback (Model 1) and

X_3 = Bottoms up intervention

Results

These data are difficult to analyze statistically: participant mortality reduces the sample size drastically if different administrations are analyzed as repeated measures on the same sample, but treating the different administrations as independent groups inflates the degrees of freedom to be used. To overcome these problems, only the posttest survey (April, 1977) data were analyzed for the 200 respondents in the pretest administration. There were no significant differences between treatments on the pretest. The central question is whether or not the interventions had an effect.

The effects of the four treatments were analyzed in April for those subjects who also participated in the January (1977) pretest. Significant differences between the treatments were found on measures of job attributes, general satisfaction, individual growth need strength, perceptions of organization effectiveness, communications, consideration, initiating structure, peer cohesion, expectations, and information sharing. Each of these achieved significance at $p \leq .05$.

The specific results indicate that subjects in the bottom up and traditional feedback (Model II) treatments were significantly more satisfied with perceptions of their jobs than the data handback treatment. Their perceptions of task significance and the amount of feedback they received from their jobs significantly improved relative to the databack and control conditions. There were, however, no significant differences between bottom up and Model II groups on any dimension of the job attributes scale.

Employee perceptions also indicate significant increases in organizational effectiveness. Model II and bottom up groups perceived themselves as producing more with greater quality. Again, these two treatments were not significantly different from each other. While the objective performance data have not been completely analyzed at this stage, the general results do not support the perceptions of improve quantity. In addition, it would appear that fluctuations in performance over time are random and nonsignificant.

A follow-up survey administered five months after the April measurement indicates that most of the significant differences dissipated. Of the variables reported here, only task significance remained significant. Several variables remained marginally significant (e.g., satisfaction with supervision, individual growth need strength, peer cohesion). In each of these instances, employees in the bottom up and Model II treatments were significantly more positive than were their data handback and no treatment control counterparts.

Finally, the moderating effects of Education, Feedback from Agents, Internal Work Motivation, Autonomy, Communication Pattern, Peer Cohesion and Leadership Style were investigated. It was found that of the seven possible covariates examined, two could be shown to serve as moderating variables: Feedback from Agents and Communication Pattern. Feedback from Agents moderated the effects on the dependent measures Specific Satisfaction and Consideration. Communication Pattern moderated the effects on the dependent measures Feedback from Agents and Consideration.

Conclusion

In general, bottom up and traditional feedback resulted in improved perception relative to the data handback and control condition. Bottom up resulted in improved affective responses to eight of the thirteen dependent variables reported here. Model II resulted in improvement in nine of the thirteen variables. Data handback resulted in few changes in either direction; communications and consideration improved while expectations and general satisfaction declined. Only the treatment control resulted in overall significant declines (on ten of the thirteen dependent variables).

The results, in terms of the comparative effectiveness of the treatments, were short-lived. Several organizational changes took place between the April survey and our September follow-up. These changes, as well as other exogenous factors, have been documented elsewhere (Lloyd, 1977). While bottom up did not outperform traditional feedback, it resulted in similar positive changes and overall did outperform data handback and no treatment controls.

REFERENCES

1. Baumgartel, H. Using employee questionnaire results for improving organizations. The survey feedback experiment. *Kansas Business Review*, 1959, *12*, 2-6.

2. Bowers, D.G. OD techniques and their results in 23 organizations: The Michigan ICL study. *Journal of Applied Behavioral Science*, 1973, *9*, 21-43.

3. Bowers, D.G. and Franklin, J.L. Survey guided development: Using human resources measurement in OD. *Journal of Contemporary Business*, 1972, 43-55.

4. Brown, L.D. Research action: Organizational feedback, understanding, and change. *Journal of Applied Behavioral Science*, 1972, *8*, 674-697.

5. Friedlander, F. and Brown, L.D. Organization development. *Annual Review of Psychology*, 1974, *25*, 313-341.

6. Hand, H.H. Estatfen, B.D. and Sims, H.P. How effective is data survey feedback as a technique of organization development? An experiment. *Journal of Applied Behavioral Science*, 1975, *11*, 333-347.

7. Kimberly, J.R. and Nielsen, W.R. Organization development and change in organization performance. *Administrative Science Quarterly*, 1975, *20*, 191-206

8. Lloyd. R.F. *Introducing Model II: A Longitudinal Investigation of the Differential Effects of A Preferred Feedback Style Over the Traditional Style on Employee Affective Reactions and Organizational Performance Measures*, Unpublished Doctoral Dissertion, Purdue University, 1977.

TABLE 2

Summary of Overall Significant Changes in April

	Feedback Variation				
Dependent Measure	Model II (1)	Bottom Up (2)	Data (3) Handback	No (4) Treatment	Neuman-Kuels Tests
Job Attributes	Improved		Declined	Declined*	1 and 2 from 3**
General Satisfaction		Declined	Declined		1 and 2 from 3 and 4*
Satisfaction/ Growth				Declined*	1, 2 and 4 from 3*
Individual Growth Need Strength			Improved	Improved	1 and 2 from 3 and 4*
Organization Effectiveness	Improved	Improved			1 and 2 from 3*
Quality	Improved	Improved		Declined*	2 from 3*
Quantity	Improved			Declined*	1 and 2 from 3 and 4*
Communications	Improved*	Improved*	Improved	Declined*	2 from 4**
Consideration	Improved*	Improved*	Improved	Declined*	2 from 4*
Initiating Structure		Improved		Declined*	2 from 4*
Peer Cohesion	Improved	Improved		Declined*	1 and 2 from 3 and 4*
Expectations	Improved*	Improved*	Declined	Declined*	1 and 2 from 3 and 4*
Information Sharing	Improved*	Improved*		Declined*	1 and 2 and 3 from 4*

* = $p \leqslant .05$
** = $p \leqslant .01$

9. Mann, F.C. Studying and creating change: A means of understanding social organization. In Bennis, W., Benne, K. and Chin, R. (Eds.) *The Planning of Change*, New York: Holt, Rinehart and Winston, 1969.

10. Miles, M., Hornstein, H., Callahan, D., Calder, P., and Schiavo, R. The consequences of survey feedback: Theory and evaluation. In Bennis, W., Benne, K., and Chin, R. (Eds.) *The Planning of Change*, New York: Holt, Rinehart and Winston, 1969, 456-468.

11. Pasmore, W.A. *Understanding organizational change: A longitudinal investigation of the effects of socio-technical system, job design, and survey feedback interventions on organizational task accomplishment and human fulfillment*. Ph.D. Thesis, Purdue University, 1976.

12. Peak, H. Attitude and motivation. In Jones, M.R. (Ed.) *Nebraska Symposium on Motivation*, 1955, 149-188.

SIMILARITIES OF MEN AND WOMEN IN PERCEPTIONS OF THEIR WORK AND THEIR AFFECTIVE RESPONSES: AN EMPIRICAL TEST

William E. Rosenbach and Valentin W. Tirman Jr.
Department of Behavioral Sciences and Leadership
U.S. Air Force Academy

ABSTRACT

Female and male workers described characteristics of their job as well as their satisfaction with the job itself and related aspects. As expected, the men and women did not differ in their perceptions of their work or their affective responses to it. The only significant difference between the two groups was their satisfaction with pay.

A growing number of studies have empirically demonstrated that supposed sex differences in personality, abilities, and attitudes about work are not valid (Lirtzman and Wahba, 1973; Reif, Newstrom, and Monczka, 1975). Decisions made about women on the sole basis of their sex, without considering individual and demographic characteristics such as background, education, experience, personality, and potential are likely to be in error. The major objective of this research was to explore the alleged differences between men and women with respect to their attitudes about their jobs and their affective responses to work.

A limited number of research studies have been conducted which inquire into the effects of gender on work-related outcomes. The major portion of the job satisfaction and motivation literature has been directed at the impact of job related attitudes on performance and affective outcomes among men. These job related attitudes have been most often dichotomized as intrinsic and extrinsic dimensions of work (Herzberg, Mausner, and Snyderman, 1959; Ewen, 1964; Friedlander, 1963, 1964; Herzberg, 1966). Herzberg and other researchers have said the intrinsic factors facilitate the personal growth and development of the individual while the extrinsic factors "maintain" the individual and contribute little to job satisfaction and performance (Malinovsky and Barry, 1965). Since the inception of the two-factory theory (Herzberg, et al., 1959), numerous studies have both supported and refuted the theory (Whitsett and Winslow, 1967; House and Wigdor, 1967).

A few research efforts have used components of the two-factory theory to study the effects of gender on work-related attitudes and affective job outcomes. Centers and Bugental (1966) showed that no consistent sex differences occurred in the overall value placed on intrinsic versus extrinsic job components. This agreed with the findings of Saleh and Lalljee (1969). They did find, however, that half of the female workers mentioned "good co-workers" as important to them, whereas only one-third of the male workers mentioned it. Further, they found that men were more concerned with the chance to use their skills in a job than were women. Finally they found that occupation was more psychologically central to men than to women, which was in basic agreement with Kuhlen's (1963) findings. Mier (1972), using a sample of females across five occupational groups, found that correlations between extrinsic needs and persistence at work were lower than the correlations between intrinsic needs and persistence at work.

In a study conducted by Ronan and Organt (1973) it was found that gender was not an important variable for determining salaried workers' satisfaction. Further, they were in agreement with Herzberg, et al., (1959) in saying that "job attitudes of women who are career oriented are more similar to those of men than are job attitudes of non-career women."

Shapiro's (1975) study showed that differences between men and women occur with respect to the relationship between job motivation, experience, preceived job responsibility, and age. His findings indicated that experience, responsibility, and age exhibited stronger positive correlations with job motivation for women than for men. A similar study by Hunt and Saul (1975) studied the relationship between age, tenure, and job satisfaction for male and female white collar workers. They found that job satisfaction was more strongly associated with age than with tenure for males, while the opposite was true for females. These findings do not contradict Shapiro's since the dependent variables were different in the two studies.

Research efforts which have attempted to compare predictions of satisfaction and job motivation for both men and women using intrinsic task attributes are disappointingly few in number. In a pioneering effort by Hackman and Lawler (1971) it was found that strong positive relationships existed between core job dimensions (variety, autonomy, task identity, and feedback from the job) and general satisfaction as well as job involvement. Wanous (1974), using Hackman and Lawler's (1971) model of work motivation found that core job dimensions all correlated strongly with general satisfaction for women having greater "higher order need strength." None of the relationships for women low on "higher order need strength" were significant.

Finally, a multinational study conducted by Kraut and Ronen (1975) demonstrated that higher order job facets were most important for predicting overall job satisfaction for men. Specific higher order job facets were job challenge, co-workers, autonomy, advancement, and skills. These findings are in rough agrreement with the conclusions drawn by Hackman and Lawler (1971), Hackman and Oldham (1974a), and Wanous (1974).

In summary, the present literature on motivation and satisfaction with respect to comparative male and female studies indicates that: (1) intrinsic factors are better predictors of affective job outcomes than extrinsic factors for both men and women, (2) certain demographic characteristics (age, tenure on the job, and education) and job characteristics (responsibility, opportunity to use skills, and interaction with co-workers) interact with sex to predict affective job outcomes, (3) individual differences (in terms of growth need strength) may moderate the relationship between job characteristics and affective work outcomes, and (4) career oriented men and women perceive job attributes in much the same way.

The present research study attempted to explore some of these relationships by comparing how male and female workers perceive the presence of core job dimensions in their job, and how they affectively responded to their work in terms of internal work motivation and general job satisfaction. In addition, specific facets of satisfaction were measured for the male and female workers. Specifically, this study hypothesized that female and male workers do not differ in their perceptions of the amount of core job dimensions present in their job or in their affective responses to their work. In addition, the authors hypothesized that women do not differ from men in their need for professional and personal growth in their job.

METHOD

The subjects were 130 members of a transportation unit of a military organization located in the southeastern United States. Thirteen of the participants in the study were women and 117 were men. All of the participants were transportation specialists and worked at the same organizational level.

The Job Diagnostic Survey (JDS) (Hackman and Oldham, 1974b) was administered to the subjects in their work setting by one of the authors. The instrument provided measures of five core job dimensions defined as follows:

1. Skill Variety - the degree to which a job requires a variety of different activities in carrying out the work, and which involves the use of different skills and talents of the worker.
2. Task Identity — the degree to which the job requires completion of a "whole" and identifiable piece of work, i.e., doing the job from beginning to end with a visible outcome.
3. Task Significance — the degree to which a job has a substantial impact on the lives or work of other people.
4. Autonomy — the degree to which the job provides substantial freedom, independence, and discretion to the worker in scheduling the work and in determining the procedure to be used in carrying it out.
5. Feedback from the Job Itself — the degree to which carrying out the work activities results in the worker obtaining direct information about the effectiveness of his or her performance.

Secondly, the JDS was used to measure two supplementary dimensions which were useful for understanding jobs and employees' reactions to them:

1. Feedback from Agents — the degree to which the employee receives clear information about his or her performance from supervisors or co-workers. This dimension is not, in a strict sense, a characteristic of the job itself; but it is included to provide information to supplement that provided by the "feedback from the job itself" dimension.
2. Dealing with Others — the degree to which the job requires the employee to work closely with other people in performing his or her job. It includes dealings with other organizational members and with external organizational clients.

Hackman and Oldham (1974a) have devised and tested a summary score which reflects the overall motivating potential of a job in terms of the core job dimensions. The motivaaing potential score (MPS) is expressed mathematically as follows:

$$\text{MPS} = \frac{\text{Skill Variety} + \text{Task Identity} + \text{Task Significance}}{3} \times \text{Autonomy} \times \text{Feedback}$$

The instrument was also used to measure two affective reactions of employees to their jobs. These reactions were viewed as personal outcomes obtained from working:

1. General Job Satisfaction — an overall measure of the degree to which the employee is satisfied and happy with the job.
2. Internal Work Motivation — the degree to which the employee is self-motivated to perform effectively on the job.

The JDS also measured separate specific facets of satisfaction with:

1. Job security.
2. Pay and other compensation.
3. Peers and co-workers (social satisfaction).
4. Supervision.
5. Opportunities for personal growth and development on the job (growth satisfaction).

Finally, the JDS was used to measure growth need strength which Hackman and Oldham (1974a) describe as a malleable individual difference characteristic which predicts how an individual will respond to a job with high motivating potential.

One-way analysis of variance (ANOVA) was performed to test for differences between men and women with respect to their perceptions of the amount of the core job dimensions present in their job, their affective responses to their job, their satisfaction with specific aspects of their job, and their growth need strength.

RESULTS

The results of the study are presented in Table 1. The descriptive statistics as well as the results of the one-way ANOVA are presented in the table.

The analysis indicated that sex did not significantly affect perceptions of any of the core or supplementary job dimensions. In addition, female and male workers' responses to their work in terms of general job satisfaction and internal work motivation were very similar. In terms of the specific facets of satisfaction, there was significant difference ($p < .02$) between men and women only for their satisfaction with pay and other compensation. The results also indicated that there was no significant difference between the female and male workers with regard to the individual difference characteristic of growth need strength.

DISCUSSION AND CONCLUSIONS

The research reported here focused on how female and male workers at the same organizational level perceive the characteristics of their job and affectively respond to their work. The hypothesis that men and women do not differ in how they perceive their job was strongly supported since the two groups did not significantly differ in their perceptions of any of the core dimensions present in their job. In addition, the hypothesis that female and male workers affectively respond to their work in a similar manner also received strong support. General job satisfaction and internal work motivation for the female and male groups were very similar. These findings support the findings of previous researchers. Specifically, Centers and Burgetal (1966) have demonstrated that little difference exists between men and women and their perceptions of intrinsic job dimensions. Rosenbach, Dailey, and Morgan (1976) have shown that organizational level affects women's perceptions of their work in the same way it affects men's perceptions. Further, the lack of difference between men and women with respect to job motivation confirms the findings of Reif, et. al. (1975).

TABLE 1

Descriptive Statistics and One-Way ANOVA of Job Dimensions and Affective Outcomes for Women and Men

Variable	Men		Women		
	M	SD	M	SD	F
Skill Variety	4.51	1.37	4.46	1.26	0.01
Task Identity	4.92	1.32	4.41	1.17	1.82
Task Significance	5.67	1.10	5.61	1.35	0.03
Autonomy	4.74	1.24	4.51	1.06	0.41
Feedback from the Job	4.98	1.24	5.05	1.53	0.04
Feedback from Agents	4.14	1.58	3.36	1.70	2.85
Dealing with Others	5.85	1.01	5.79	1.00	0.04
MPS	126.31	61.28	118.20	71.17	0.20
General Job Satisfaction	4.24	1.34	4.08	0.87	0.17
Internal Work Motivation	5.12	0.95	5.37	1.01	0.82
Pay Satisfaction	4.33	1.64	5.42	1.27	5.37*
Security Satisfaction	4.74	1.37	5.04	1.05	0.58
Social Satisfaction	5.36	1.03	5.21	1.55	0.24
Supervisory Satisfaction	4.86	1.51	4.36	1.85	1.25
Growth Satisfaction	4.73	1.36	4.37	1.41	0.82
Growth Need Strength	5.55	1.30	6.19	1.06	2.93

*$p < .02$

The analysis of the specific facets of satisfaction indicated that the only significant difference between the female and male workers was for satisfaction with pay and other compensation ($p. < .02$). This difference can be explained if one considers that in the focal organization, men and women are equally compensated for performing the same job. This equality of pay without regard to sex does not exist in all organizations, especially in those located in the section of the country where the reported study was conducted. The women probably perceived the inequity in compensation that existed for women in other organizations and were therefore more satisfied with their pay. In addition, job opportunities in the local community were limited, especially for women. The reported findings are supported by Reif, Newstrom and St. Louis (1976) who also found that in the lower organizational level, women were significantly more satistied with pay and other economic rewards than were men.

The similarity of the female and male workers in terms of the individual difference variable, growth need strength, was expected and confirms previously reported research (Hackman and Oldham, 1974a). The implication is that growth need strength moderates the way an individual reacts to a job with high motivating potential without regard to sex. Or in other words, men and women are similar in their need to grow personally and professionally in their work.

This study was limited by the small sample of female workers; however, it does adequately demonstrate that female workers are more similar to men than different in terms of how they perceive their work and react to it. Rather than continuing to explore the alleged differences between female and male workers, future research should concentrate on improving the quality of work life for all workers, regardless of sex.

REFERENCES

1. Armstrong. T.R. "Job Content and Context Factors Related to Satisfaction for Different Occupational Levels," *Journal of Applied Psychology*, 1971, *55*, 57-65.

2. Centers, R. and D.E. Bugental. "Intrinsic and Extrinsic Job Motivations Among Different Segments of the Working Population," *Journal of Applied Psychology*, 1966, *50*, 193-197.

3. Ewen, R.B. "Some Determinants of Job Satisfaction: A Study of the Generality of Herzberg's Theory," *Journal of Applied Psychology*, 1964, *48*, 161-163.

4. Friedlander, F. "Underlying Sources of Job Satisfaction," *Journal of Applied Psychology*, 1963, *47*, 246-250.

5. ______________. "Job Characteristics as Satisfiers and Dissatisfiers," *Journal of Applied Psychology*, 1964, *48*, 388-392.

6. Hackman, J.R. and E.E. Lawler. "Employee Reactions to Job Characteristics," *Journal of Applied Psychology Monograph*, 1971, *55*, 259-285.

7. Hackman, J.R. and G.R. Oldham. *Motivation Through the Design of Work: Test of a Theory*. Tech. Rpt. No. 6, Department of Administrative Sciences, Yale University, 1974(a).

8. ______________. *The Job Diagnostic Survey: An Instrument for the Diagnosis of Jobs and the Evaluation of Job Redesign Projects*, Tech. Rpt. No. 4, Department of Administrative Sciences, Yale University, 1974(b).

9. Herzberg, F. *Work and the Nature of Man*. Cleveland: World Publishing, 1966.

10. Herzberg, F., B. Mausner, and B. Snyderman. *The Motivation to Work*. New York: John Wiley and Sons, 1959.

11. House, R.J. and L.A. Wigdor. "Herzberg's Dual Factor Theory of Job Satisfaction and Motivation: A review of the Evidence and a Criticism," *Personnel Psychology*, 1967, *20*, 369-389.

12. Hrebiniak, L.G. "Effects of Job Level and Participation on Employee Attitudes and Perception of Influence," *Academy of Management Journal*, 1974, *17*, 649-662.

13. Hunt, J.W. and P.N. Saul. "The Relationship of Age, Tenure, and Job Satisfaction in Males and Females," *Academy of Management Journal*, 1975, *18*, 690-701.

14. Kraut, A.I. and S. Ronen. "Validity of Job Facet Importance: A Multinational Multicriteria Study," *Journal of Applied Psychology*, 1975, *60*, 671-677.

15. Kuhlen, R.G. "Needs, Perceived Need Satisfaction Opportunities, and Satisfaction with Occupation," *Journal of Applied Psychology*, 1963, *47*, 56-64.

16. Lirtzman, S.I. and M.A. Wahba. "A Managerial Myth: Differences in Coalitional Behavior of Men and Women in Organizations," *Proceedings of the 32nd Annual Meeting of the Academy of Management*, Academy of Management, 1973.

17. Malinovsky, M. and J. Barry. "Determinants of Work Attitudes," *Journal of Applied Psychology*, 1965, *49*, 446-451.

18. Mier, E.I. "Relationship Between Intrinsic Needs and Women's Persistence at Work," *Journal of Applied Psychology*, 1972, *56*, 293-296.

19. Reif, W.E., J.W. Newstrom and R.D. St. Louis. "Sex as a Discriminating Variable in Organizational Reward Decisions," *Academy of Management Journal*, 1976, *19*, 469-476.

20. Reif, W.E., J.W. Newstrom, and R.M. Monczka. "Exploding Some Myths About Women Managers," *California Management Review*, 1975, *17*, 72-79.

21. Rosenback, W.E., R.C. Dailey and C.P. Morgan. "Differences Among Women in Perceptions of Their Jobs," *Academy of Management Proceedings*, 1976, 472-476.

22. Ronan, W.W. and G.J. Organt. "Determinants of Pay and Pay Satisfaciion," *Personnel Psychology*, 1973, *26*, 502-520.

23. Saleh, S.D. and M. Lalljee. "Sex and Job Orientation," *Personnel Psychology*, 1969, *22*, 465-471.

24. Shapiro, H.J. "Job Motivations of Males and Females: An Empirical Study," *Psychological Reports*, 1975, *36*, 647-654.

25. Wanous, J.D. "Individual Differences and Reactions to Job Charateristics," *Journal of Applied Psychology*, 1974, *59*, 616-622.

26. Whitsett, D.A. and E.K. Winslow. "An Analysis of Studies Critical of the Motivator-Hygiene Theory," *Personnel Psychology*, 1967, *20*, 391-415.

BRAIN FUNCTIONAL SPECIALIZATION AND FLYING PERFORMANCE: A PROPOSED STUDY*

*Kirmach Natani***

Clinical Sciences Division
and
Crew Technology Division
School of Aerospace Medicine
Brooks AFB, Texas 78235

ABSTRACT

It is proposed that functional specialization in the brain may provide an index of pilot information processing and decision-making skills. Functional specialization will be assessed by computer-administered lateralized optional shift task. Link GAT-1 simulator performance will be computer scored. A positive correlation between lateralization and flying skills is predicted.

INTRODUCTION

Cognitive neuropsychology is concerned with the role individual brain systems play in the organization and direction of complex human activities. Each cerebral hemisphere in man apparently constitutes an integrated system possessing highly specialized functions. It is now generally accepted that in most individuals the left hemisphere performs language related sequential, logical processes under conscious control while the right hemisphere controls parallel, spatially oriented semi-conscious processes (Milner, 1974). Behavioral consequences of interaction between these contrasting specializations have been of relatively recent interest. Earlier work emphasized the mental functions (Diamond, Balvin & Diamond, 1963; Simons, Flinn & Hartman, 1963). Recent reports have included laterality factors and have also suggested that cerebral dominance is a life span developmental process (Brown & Jaffe, 1975; Luria, 1975). The importance of interhemispheric relationships as expressed, for example, by Luria and Simernitskaya (1977) has contributed significantly to the theoretical basis for this study.

PROBLEM

The pilot's role includes decisive control of his aircraft as well as the swift and expeditious solution of unforeseen emergencies or enemy engagements for the military aviator. The two tasks require both precise psychomotor and high level cognitive skills. The laterality literature suggests the hypothesis that with training and experience these two types of specialized skills may become functionally localized in the two hemispheres of the brain. These skills can also be dichotomized as automatic (right hemisphere) and voluntary (left hemisphere) functions. Such an arrangement would contribute substantially to the efficient simultaneous performance of both functions when required. A corollary to this hypothesis is that an individual possessing a brain already functionally organized for the efficient performance of tasks in both hemispheres should be able to learn tasks requiring simultaneous use of both sides of the brain at an accelerated rate. On the other hand, poorly organized cerebral lateralization would suggest the need to compensate in some way for the lack of functional structure. Compensation should be evident in slower responses on individual task trials and in the need for extended practice in learning complex new tasks.

The purpose of this study is to investigate these possibilities by obtaining and comparing simulator learning curves from individuals selected for indications of well and poorly established functional symmetry for perceptual learning with a task believed to provide a measure of cerebral lateralization (Natani & Parsons, Note 1).

METHOD

The selection task to be used combines a well known experimental paradigm in concept learning with generally accepted techniques for demonstrating laterality effects in the normal adult human brain. The learning task is the optional shift discrimination reversal paradigm developed by the Kendlers (H.H. & T.S. Kendler, 1975). The subject is required to determine the classification scheme selected by the experimenter for a set of multidimensional stimuli presented monocularly in the contralateral visual field. The task has been described in detail by Natani (1977). In this study all stimulus presentation, feedback, and data collection operations will be performed by a DEC PDP-12 computer.

A Pilot Performance Evaluation System (Henry, Note 2; Henry, Turner & Matthie, Note 3) developed at the USAF School of Aerospace Medicine and coded for PDP-12 administration (Perelli, Note 4) will provide performance data based on error scores for six instruments in a Singer-Link GAT-1 flight simulator during a standardized flight maneuver test. Parameters to be used in evaluating flying performance are (1) altitude ± 40 ft, (2) heading ± 2½°, (3) airspeed ± 5 mph, (4) vertical velocity ± 150 fpm, (5) turn rate ± 0.5° sec, and ball angle ±5° vertical. Professional pilot/instructor raters have indicated that the simulator task provides a reasonably good evaluation of flying skills required in T-37 or T-38 aircraft flown on training missions.

---

*Prepared for presentation at the Sixth Psychology in the DoD Symposium, US Air Force Academy, Colorado 80840, April 20-22, 1978.

**National Research Council Postdoctoral Resident Research Associate in Neuropsychology. Bryce O. Hartman, Ph.D., Scientific Advisor.

Volunteer subjects will be Airman Basic graduates from Lackland AFB. During the study these individuals will reside at Brooks AFB and also participate in an ongoing extensive clinical neuropsychological testing program. A randomized between subjects experimental design will be used to statistically compare the flight performance scores of those individuals identified as (1) fast learning-fast shifters, (2) fast learning-slow shifters, and (3) slow learning-fast shifters by the laterality selection task. There will be a minimum of eight subjects in each of the three groups. Voluntary consent forms will be obtained from all the participants in accordance with AFR 80-33.

RESULTS

Preliminary results with the use of the lateralized optional shift task indicate that the performance of one hemisphere, in terms of total trials to criterion, can be used to predict performance in the opposite hemisphere. This finding suggests that the task may provide a good measure of bilateral symmetry for perceptual learning. However, for this particular task an analysis of errors, rather than trials to criterion, provides the best indication of lateral differences in information processing strategies or functional specialization. Proficient individuals have also provided test scores indicating a pattern of intelligence, cognitive style and personality traits which may be associated with well established functional organization (Levy & Natani, Note 5).

The Pilot Evaluation System has been standardized to the extent that cutoff scores are available for various levels of training and flying proficiency. It is expected that those subjects showing indications of good bilateral symmetry will also produce the best flying performance scores.

DISCUSSION

The advantages of this approach appear to be: (1) a neurophysiologically oriented theoretical basis for predicting both performance advantages and decrements, (2) the availability of a highly flexible selection task which can be easily modified in terms of complexity and modality of presentation as future needs may require and (3) the use of a sophisticated performance measure which has been rated by professional pilots as doing "moderately well" in simulating the actual behavioral demands of flying training aircraft.

Implications of theory also suggest the possibility of extending this approach to the development of remedial training procedures and evaluation of individual potential for performance decrement under stress. It should be noted that a hemisphere specific task will not be used to determine laterality. At this time it is difficult to assume that enough is known about the nature of cerebral functional specialization to choose a supposed right or left hemisphere task that will have high validity for all individuals that might be tested. Tasks oriented toward the function of a given hemisphere may also obscure the presence of a strong hemispheric bias or condition of 'functional hemisphericity' (Bogen, 1969; Note 6). Such a condition in an individual may actually interfere with the bilateral acquisition of complex new skills. Therefore, the primary selecting criterion will be essentially equivalent performance in both hemispheres with the same type of task. It is expected that symmetrical performance will indicate well integrated intrahemispheric functional structures while it is assumed that the two hemispheres will not solve the task with the same processing strategies. Measurement of bilateral processing performance appears to be more important at this stage of knowledge about the mechanisms of cerebral function and interhemispheric interactions.

A change in task conditions has been used when testing the second hemisphere to control for practice and transfer of training effects. For experimental purposes the procedure has been to counterbalance for hemisphere and task conditions across subjects. In the event that such a procedure is ever developed into a psychometric tool then a standardized method of presentation for individual subjects will need to be established.

REFERENCE NOTES

1. Natani, K. & Parsons, O. A. "Tachistoscopic discrimination reversal shift learning: an index of cerebral lateralization?" Paper in preparation, 1978.

2. Henry, P. H. "Subject instruction manual for the pilot performance evaluation system" (Report SAM-TR-74-40). San Antonio, Texas, USAF School of Aerospace Medicine, Brooks AFB, October 1974.

3. Henry, P. H., Turner, R. A. & Matthie, R. B. "An automated system to assess pilot performance in a Link GAT-1 trainer" (Report SAM-TR-74-41). San Antonio, Texas, USAF School of Aerospace Medicine, Brooks AFB, October 1974.

4. Perelli, L. P. "Effects of fatigue stressors on psychobiological indices of air crew performance. Unpublished experimental protocol, USAF School of Aerospace Medicine, 1977.

5. Levy, R. A. & Natani, K. "Adverse reactions in biofeedback therapy: implications for a theory of functional hemisphericity." Paper in preparation, 1978.

6. Bogen, J. E. "Personal communication," annual meeting of International Neuropsychology Society, Santa Fe, New Mexico, February, 1977.

REFERENCES

1. Bogen, J. E. The other side of the brain II: an appositional mind. *Bulletin of the Los Angeles Neurological Society*, 1969, 34:135-162.

2. Brown, J. W. & Jaffe, J. Hypothesis on cerebral dominance. *Neuropsychologica*, 1975, 13:107-110.

3. Diamond, S., Balvin, R.S. & Diamond, F.R. *Inhibition and Choice: A Neurobehavioral Approach to Problems of Plasticity in Behavior*. New York: Harper & Row, 1963.

4. Kendler, H. H. & Kendler, T. S. From discrimination learning to cognitive development: a neobehavioristic odyssey. In W. K. Estes (Ed.) *Handbook of Learning and Cognitive Processes* (Vol. I). Hillsdale, N.J.: Lawrence Erlbaum Associates, 1975.

5. Luria, A. R. Neuropsychology: Its sources, principles and prospects. In F. G. Worden, J. P. Swazey & G. Adelman (Eds.) *The Neurosciences: Paths of Discovery*. Cambridge: MIT Press, 1975.

6. Luria, A. R. & Simernitskaya, E. G. Interhemispheric relations and the functions of the minor hemisphere. *Neuropsychologia*, 1977, *15*:175-178.

7. Milner, B. Hemispheric specialization: scope and limits. In F. O. Schmitt and F. G. Worden (Eds.) *Neurosciences Third Study Program*. Cambridge: MIT Press, 1974.

8. Natani, K. Laterality effects in a tachistoscopic optional shift task in young adults (Doctoral dissertation, University of Oklahoma Health Sciences Center, 1977). *Dissertation Abstracts International*, 1977, *38*, (University Microfilms No. 77-16, 744).

9. Simons, D., Flinn, D. & Hartman, B. Psychophysiology of high altitude experience. In N. M. Burns, R. M. Chambers and E. Hendler (Eds.) *Unusual Environments and Human Behavior*. New York: Free Press of Glenco, 1963.

MATCHING INITIAL PERFORMANCE AND THE MEASUREMENT OF SEX DIFFERENCES

Beverly H. Williges, Robert C. Williges, and Rickey E. Savage

Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Women are increasingly being selected to perform traditionally male jobs, including tasks requiring a high degree of motor skill. However, little data exist on female motor performance and the effects of training on the development of motor skills. Two studies that examined sex differences in a two-dimensional pursuit tracking task are presented. Results of these studies imply the need for a multifactor matching procedure to study sex differences in motor skills.

Increasingly the "man" in many man-machine systems is a woman. For example, women are being used in vehicular control systems requiring a high degree of motor skill. And, yet, little data exist on female motor skill performance and effects of various training models on the development of motor skill. A recent survey by Hudgens and Billingsley (1978) analyzed the sex of subjects included in studies published in *Human Factors* and *Ergonomics* for the time period of 1965 through 1976. Their results indicate that only a quarter of the studies included females either exclusively (6%) or with males (19 %). In addition, the percentage of studies that include women, reported in these two journals, has not increased in recent years. Human factors data on women are urgently needed in order to establish appropriate training programs for women and to redesign equipment and jobs for women, when necessary.

SEX DIFFERENCES IN MOTOR SKILLS

Evidence for differences in tracking skill favoring males has been obtained with children (Ammons, Alprin, and Ammons, 1955) and with college students (Noble, 1970). However, it is not safe to conclude that males and females, therefore, differ in motor abilities or learning rates. Singer (1975) suggests that these sex differences may be the result of transfer of non-specific training, previous experience with the task, motivational differences, and/or sociocultural factors such as a need for social approval.

Prior Experience

Williges and Williges (1977) reported the results of a study to examine sex differences in learning a two-dimensional pursuit tracking task using a fixed difficulty training model, an adaptive training model (Kelley, 1969), or a learner-centered model. They reported a highly reliable ($p < .001$) sex difference in training time-to-exit favoring males. However, no reliable differences in tracking error between males and females were found in transfer ($p = .48$) even during those periods during the transfer task when the level of tracking difficulty exceeded the maximum level of task difficulty during training. These results suggest that, although women may require some additional training time initially on motor-control tasks similar to the one used in this study, no performance differences by sex should be expected once training is accomplished.

Possibly many of the sex differences in motor skills are more correctly measures of prior experience with similar motor-control tasks. In the Williges and Williges (1977) study, an analysis of variance on vector error for the fixed difficulty training group in the first 3 minutes of training was conducted and revealed a reliable sex difference, $F(1, 10) = 9.45$, $p < .05$. At least in the fixed-difficulty condition, males and females were not equated in terms of initial motor performance.

MATCHING PROCEDURES TO STUDY SEX DIFFERENCES

One disadvantage of a between-subjects design is that quite by accident subjects randomly assigned to one or more treatment group can have significantly different prior experience, skill, or motivation than subjects assigned to the remaining treatment groups. Another possibility is that of drawing from different populations in terms of initial skill level. When subjects are drawn from two distributions which overlap but have different means, the most likely result, given random sampling, is two samples with different means (unmatched).

However, in training situations a within-subject design is impossible because of the baseline problem. To make the subject population as uniform as possible, a common procedure is to match subjects on some concomitant variable that is correlated with the dependent variable. Usually a matching variable is selected based either upon previous documentation of its appropriateness or preliminary experimentation

Selecting a Matching Variable

To minimize initial performance differences between males and females, the present authors chose to match male and female subjects for a second study of sex differences in pursuit tracking. In order to select an appropriate matching variable, a pretest was conducted using four simple motor tasks: pursuit rotor tracking, a one-dimensional (y-axis) pursuit tracking task, a one-dimensional (x-axis) compensatory tracking task in which task difficulty increased over time, and a television game (handball) in which the object was to maximize time in the game. Ten male subjects were pretested, and then four received fixed-difficulty training and six received adaptive training. Table 1 summarizes the Pearson Product Moment Correlations of each pretest with training time-to-exit and with transfer tracking error. The only pretest which resulted in a consistently high correlation across the two training conditions was the rotary pursuit.

Subsequently, 90 subjects were tested with the pursuit rotor in order that 30 male and 30 female subjects could be matched and assigned to three training conditions. Table 2 provides means and standard deviations by sex for the original random sample of subjects and for the matched sample. Obviously, matching did reduce the performance differences by sex on the pursuit rotor task. Methodological details

of this study are given in Williges, Williges, and Savage (1977). Of importance here is the adequacy of the matching procedure employed. Table 3 provides a comparison of Pearson Product Moment Correlations in the preliminary and main studies for training and transfer. In general the correlations of pursuit rotor scores with training time dropped in the main study particularly for females in the adaptive training condition. Obviously, a serious flaw in the preliminary study was the failure to use any female subjects. However, even for male subjects the correlations dropped substantially, perhaps to the point of rendering the pursuit rotor scores useless as a matching variable. It becomes apparent that matching male and female students on initial motor skill performance is not an easy task. What procedure, then, can be used?

TABLE 1

Pearson Product Moment Correlations of Pretest Tasks with Training Time-to-Exit and with Transfer Tracking Error

Training Condition	Pretest PR	PT	CT	TV
	Training Time-to-Exit			
Adaptive	-.762	-.292	.164	-.249
Fixed	-.711	.886	-.496	.352
Overall	-.550	.147	-.138	-.065
	Transfer Tracking Accuracy			
Adaptive	-.471	-.365	.051	-.642
Fixed	-.608	.633	-.827	.422
Overall	-.425	.106	-.447	-.340

PR = pursuit rotor tracking
PT = y-axis pursuit tracking
CT = x-axis compensatory tracking
TV = television game

TABLE 2

Means and Standard Deviations of Pursuit Rotor Scores by Sex for Random and Matched Samples of Subjects

Sex	Random Sample N	$\bar{X}$	σ	Matched Sample N	$\bar{X}$	σ
Male	51	8.92	4.90	30	6.14	2.79
Female	39	4.95	2.99	30	5.61	2.92

Multi-Factor Matching

As part of the pretest package five tests in addition to the pursuit rotor were administered. These tests were various measures of information processing capacity of the student and are described in more detail by Savage, Williges, and Williges (1978). Standardized scores from all six tests were used to generate least squares regression equations for each sex and training condition combination. See Table 2 for the multiple R^2, estimate of shrinkage, R_s^2 (Kerlinger and Pedhazur, 1973), and number of subjects in each sex training condition combination. For estimating training time-to-exit a substantial improvement was obtained using a multi-factor predictor. For estimating transfer tracking error, the multi-factor predictor was useful only for females.

TABLE 3

Pearson Product Moment Correlations of Pursuit Rotor Scores, Multiple R^2, Estimate of Shrinkage (R_s^2), and Number of Subjects for Training and Transfer

	Preliminary Study		Main Study			
	N	r	N	r	R^2	R_s^2
Training Time-to-Exit						
Adaptive Training						
Male	6	-.762	18	-.600	.750	.717
Female	*	*	13	-.319	.871	.806
Fixed Training						
Male	4	-.711	16	-.446	.803	.758
Female	*	*	12	-.454	.476	.364
Transfer Tracking Error						
Adaptive Training						
Male	6	-.471	18	-.215	.545	.484
Female	*	*	13	-.193	.906	.875
Fixed Training						
Male	4	-.608	16	-.307	**	**
Female	*	*	12	-.690	.807	.697

* No females in the preliminary study.
** No reliable regression equation.

Practice and Task-Specific versus General Abilities

Various researchers (Adams, 1957; Fleishman, 1960; Jones, 1972) have noted that in complex tasks initial task performance is often related to general abilities. As training proceeds, however, task-specific variables become more important predictors.

In the test battery administered, all the tests besides the pursuit rotor measure cognitive/perceptual (general) abilities. One could speculate that the diminished ability of the test battery to predict transfer performance for males reflects a movement to a more advanced stage of training where task-specific variables have more importance as predictors.

IMPLICATIONS

It is obvious that a data base on female performance is essential to facilitate the incorporation of women into traditionally

all-male systems. However, acquiring a useful data base on women is not a simple task. Results of the two studies presented in this paper imply a need for more sophisticated multi-factor procedures for matching subjects to study sex differences in motor skills.

ACKNOWLEDGMENTS

Contractual support for this project was provided by the Life Sciences Program, Air Force Office of Scientific Research, Grant Number 77-3161. Captain Jack A. Thorpe served as the scientific monitor.

REFERENCES

Adams, J.A. The relationship between certain measures of ability and the acquisition of a psychomotor criterion response. *The Journal of General Psychology*, 1957, *56*, 121-134.

Ammons, R.B., Alprin, S.I., and Ammons, C.H. Rotary pursuit performance as related to sex and age of pre-adult subjects. *Journal of Experimental Psychology*, 1955, *49*, 127-133.

Fleishman, E.A. Abilities at different stages of practice in rotary pursuit performance. *Journal of Experimental Psychology*, 1960, *60*, 162-171.

Hudgens, G.A. and Billingsley, P.A. Sex: The missing variable in human factors research. *Human Factors*, 1978, in press.

Jones, M.B. Individual differences. In R.N. Singer (Ed.) *The psychomotor domain: Movement behavior*. Philadelphia: Lea & Febiger, 1972.

Kelley, C.R. The measurement of tracking proficiency. *Human Factors*, 1969, *11*, 43-64.

Kerlinger, F.N. and Pedhazar, E.J. *Multiple regression in behavioral research*. New York: Holt, Rinehart, and Winston, 1973.

Noble, C.E. Acquisition of pursuit tracking skill under extended training as a joint function of sex and initial ability. *Journal of Experimental Psychology*, 1970, *86*, 360-373.

Savage, R.E., Williges, R.C., and Williges, B.H. Individual differences in motor skill training. *Proceedings of the Sixth Psychology in the DoD Symposium*, 1978.

Singer, R.N. *Motor learning and human performance*. New York: Macmillan, 1975.

Williges, B.H. and Williges, R.C. Learner-centered versus automatic adaptive motor skill training. *Journal of Motor Behavior*, 1977, 9, 325-331.

Williges, B.H., Williges, R.C., and Savage, R.E. Models for automated motor skills training. *Proceedings of the 21st annual meeting of the Human Factors Societyt*, 1977, 18-22.

INDIVIDUAL DIFFERENCES IN MOTOR SKILL TRAINING

Ricky E. Savage, Robert C. Williges, and Beverly H. Williges

Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Fifty-one males and 39 females were given five paper-and-pencil tests (spatial scanning, visual memory, perceptual speed, spatial orientation, and perceptual style) and a motor skill test to develop regression equations that predict training time for a two-dimensional pursuit tracking task. One of three training conditions (fixed difficulty, adaptive, shifting difficulty) was randomly selected for each subject. Regression equations for the different training conditions accounted for as much as 87% of the variance in training time. Results are discussed in terms of predicting an optimal training condition for the individual.

Recent interest in the relationship between information processing skills and motor skills may have implications for investigating individual differences in motor learning. Marteniuk (1976) states that there are three basic information processing mechanisms involved in any motor skill: the perceptual, decision, and effector mechanism. Much of the emphasis in his discussion of information processing in motor skills is that limitations in any of these mechanisms can limit motor performance. If these perceptual or cognitive processes can be measured to determine differences in capacities, then individual differences can be studied as they relate to motor skills. The purpose of this study was to use a battery of test scores to generate regression equations in order to predict training time and transfer performance in a pursuit tracking task. With these equations it might be possible to assign an individual to an optimal training condition based on his or her information processing characteristics.

METHOD

Subjects

Fifty-one male and 39 female undergraduate volunteers participated in the experiment. Subjects were right-handed, naive to the experimental task, and paid for their participation.

Tests

A pretest battery of six tests was given to each subject. The first test measured time-on-target across six, 30-s trials on a pursuit rotor moving at 60 revolutions per minute. The second test was the Embedded Figures Test (Watkin, Oltman, Raskin, and Karp, 1971) which measures the perceptual ability of field independence and field dependence. The remaining four tests were chosen from a paper-and-pencil battery developed by Ekstrom, French, Harman, and Dermen (1976) which was designed to measure various cognitive components related to information processing. These tests included: the Map Memory Test for a measure of visual memory; the Cube Comparison Test for a measure of spatial orientation; the Identical Pictures Test for a measure of perceptual speed; and the Maze Tracing Test for a measure of spatial scanning.

Training Task and Procedures

The training and transfer tasks used a Digital Equipment Corporation PDP11/10 digital computer linked to a CRT display and a two-axis pressure control stick. The tracking task involved two random band-limited functions which determined the X-Y coordinates of the forcing function symbol ("X") on the display; the control output ("O") was generated from the isometric controller. Feedback bars were used to provide information on error and level of difficulty. The control system was an 80% acceleration system where the maximum gain output of the control stick was 50.8 cm/s. Task difficulty was manipulated in terms of maximum movement of the forcing function symbol. The criterion level of task difficulty during training was 20.3 cm/s. The error tolerance was 10% of the screen diagonal (18 cm by 18 cm).

Following completion of the six tests each subject learned a pursuit tracking task using one of three possible training approaches: fixed difficulty, shifting difficulty, or adaptive. The fixed difficulty condition is the traditional approach to motor skill training in which trainees are presented the criterion task initially and their error decreases as training progresses. A more individualized approach is the adaptive training condition which involves a set of decision rules to manipulate task difficulty in a closed-loop system (Kelly, 1969). Student performance is monitored and compared to a standard by adjusting the difficulty of the task so that the performance of the student is relatively stable throughout training. The shifting difficulty condition evolved from an earlier study (Williges and Williges, 1977) in which subjects controlled the level of task difficulty. Many subjects initially set the task difficulty at a low level and later shifted the difficulty to the criterion level. In the shifting difficulty condition, the level of difficulty during Trial 1 was at a low level (0.2 cm/s); in Trial 2 and all sequent trials, the task difficulty was at the criterion level.

Transfer Task

A 5-minute rest occurred between the training and transfer tasks. The 7-minute transfer task was similar to the training task except no performance information was provided. Three levels of task difficulty in terms of the maximum speed of the forcing function were presented. The three levels of difficulty were the exit criterion (20.3 cm/s), more difficult than the exit criterion (30.5 cm/s), and less difficult than the exit criterion (10.2 cm/s).

RESULTS AND DISCUSSION

Scores from the six tests were used to generate least squares regression equations for each of the three training conditions. Time-to-exit in training and mean performance during the *first six minutes in transfer* were used as dependent measures. Separate equations were also calculated for males and females in each training condition. The best regression equations resulting from five stepwise procedures were chosen which could account for the most variance with the least variables. Each of the resulting regression equations used standarized (Z) scores for the predictor pretest variables.

TABLE 1

The Multiple R^2, Estimate of Shrinkage, R_s^2, and Number of Subjects in Each Regression Equation.

	Adaptive			Fixed			Shift		
	R^2	R_S^2	N	R^2	R_S^2	N	R^2	R_S^2	N
Overall	.717	.673	31	.639	.610	28	.566	.518	31
Male	.750	.717	18	.803	.758	16	.614	.485	17
Female	.871	.806	13	.476	.364	12	.731	.650	14

Table 1 presents the multiple R^2, the estimate of shrinkage, R_s^2, (Kerlinger and Pedhazur, 1973), and the number of subjects for each equation in the three training conditions using time-to-exit as the dependent variable. Each of these regression equations was statistically significant ($p < .05$), and the shrinkage was minimal with the exception of the shift-male subjects and the fixed-female subjects. It is important to note that when the regression equations in Table 1 are broken down by sex, a larger proportion of variance can be accounted for with the only exception being the equation for the *fixed female subjects*. This observation would seem to indicate that assigning subjects to an optimal training condition (based on predicted time-to-exit scores) should consider the sex of the subject.

TABLE 2

Regression Equations for Training Time-to-Exit by Training Condition and Sex

Adaptive/Male TE = 1069.50 + 509.86 EF + (-325.06) MM
Adaptive/Female TE = 1967.61 + (-492.83) MM + 391.02 IP + (0358.42) CC + 196.43 PR
Fixed Difficulty/Male TE = 867.84 + 412.35 IP + 333.12 EF
Fixed Difficulty/Female TE = 1551.25 + 467.82 EF + 144.51 MM
Shifting Difficulty/Male TE = 857.38 + (-361.85) EF + (-354.74) CC + (-335.53) PR + (-177.03) MT
Shifting Difficulty/Female TE = 1219.86 + (-266.34) PR + (-223.36) CC + (-199.44) MT

TE – Training Time-to-Exit
PR – Pursuit Rotor (continuous motor skills)
EF – Embedded Figures Test (field *independence*)
MM – Map Memory Test (visual memory)
IP – Identical Pictures Test (perceptual speed)
CC – Cube Comparison Test (spatial orientation)
MT – Maze Tracing Test (spatial scanning)

Regression equations were also generated in the same manner using the mean transfer performance as the dependent variable. These equations were generally either unreliable or accounted for only a small proportion of the variance. The exceptions were the regression equations for the adaptive training condition which accounted for as much as 90% of the variance.

In determining which regression equations should be used to predict time-to-exit in each training condition in order to assign an individual to an optimal training strategy, it is apparent that the equations for the transfer task are too unreliable. On the other hand, the equations based on sex for predicting time-to-exit appear to be the most reliable. These six regression equations are presented in Table 2. It should be noted that across the various equations by sex and training condition, different sets of predictor variables are used in each equation, not just different weightings. This observation also argues for the need for regression equations based on the sex of the individual.

CONCLUSION

Although the results of this study demonstrate that it is possible to predict reliably training time-to-exit when individual differences due to the sex of the subject are considered, additional research needs to be directed toward cross-validating these equations with a new sample of subjects. Future research is also needed to validate the potential use of these equations in placing individuals in an optimal training condition. This would involve matching subjects to an optimal training condition based on their predicted scores and comparing their training effectiveness to mismatched subjects. If the prediction equations can withstand this experimental scrutiny, then meaningful implications can be stated concerning the role of individual differences in choosing the most appropriate training strategy.

ACKNOWLEDGMENTS

Contractual support for this project was provided by the Life Sciences Progam, Air Force Office of Scientific Research, Grant Number 77-3161. Captain Jack A. Thorpe served as the scientific monitor. The authors wish to thank John E. Evans for software support.

REFERENCES

Ekstrom, R.B., French, J.W., Harman, H.H., and Dermen, *Manual for kit of factor-referenced cognitive tests*. Princeton, New Jersey: Educational Testing Service, 1976.

Kelly, C.R. What is adaptive training? *Human Factors*, 1969, *11*, 547-556.

Kerlinger, F.N. and Pedhazur, E.J. *Multiple regression in behavioral research*. New York: Holt, Rinehart, and Winston, 1973.

Marteniuk, R.G. *Information processing in motor skills*. New York: Holt, Rinehart, and Winston, 1976.

Williges, B.H. and Williges, R.C. Learner-centered versus automatic adaptive motor skill training. *Journal of Motor Behavior*, 1977, *9*, 325-331.

Witkin, H.A., Oltman, P.N., Raskin, E., and Karp, S.A. *A manual for the embedded figures test*. Palo Alto, California: Consulting Psychologists Press, 1971.

SIMULATOR TRAINING REQUIREMENTS AND EFFECTIVENESS STUDY A PROGRESS REPORT

Bertram W. Cream

Air Force Human Resources Laboratory
Advanced Systems Division, Wright-Patterson AFB Ohio

INTRODUCTION

This paper will describe and provide a progress report on an ongoing research program entitled "Simulator Training Requirements and Effectiveness Study (STRES)." The basic objectives of this program are to systematically evaluate the effectiveness of training, cost of ownership, advantages, disadvantages and constraints of major aspects of simulator training in relation to accomplishment of specific training requirements. To meet this objective, it will be necessary to define, describe, collect, analyze, document and present relevant information bearing upon the cost and training effectiveness of important characteristics of flight simulators and other ground based training devices. With the past emphasis on providing necessary training and sustaining operational readiness using simulators as part of a total training program, the majority of R&D funding has gone towards development and improvements in simulation engineering technology. As a consequence, there have been numerous advances in such technology; however, significantly less attention has been directed towards the ways these advances could be best used by operational personnel to support their training programs. As a consequence, the current data base is insufficient to allow making tradeoff decisions concerning training, effectiveness and costs. Also obvious is the need to provide pragmatic guidance on techniques for maximizing the effectiveness and efficiency of devices currently in inventory and on order.

OBJECTIVES

The basic objectives of this proggam are to define, collect, analyze and present data, findings and conclusions relevant to the issues stated above. These issues have been structured around four broadly stated objectives:

(1) Develop criteria for matching training requirements with simulator device features and fidelity considerations in order to effectively and efficiently accomplish specified training requirements.

(2) Define principles of effective and efficient utilization of training devices as one part of a larger training system.

(3) Develop criteria for matching instructional features and techniques with the achievement of specific training requirements.

(4) Collect data and develop models of the factors that influence the cost and "worth-of-ownership" of these training devices.

Because these objectives are so broad, some method was necessary to focus the study on those issues of primary concern. Towards this end, eight "high value" tasks (HVT) were identified. The "high value tasks" are as follows: (1) individual and formation landings and take-off; (2) close formation flight and trail formation, both close and extended; (3) aerobatics; (4) spin, stall and unusual attitude recognition, recovery and prevention; (5) low level terrain following flight; (6) air refueling; (7) air-to-air combat; and (8) air-to-ground weapons delivery. These tasks were identified as "high value" because the current capability of simulation devices for training these tasks is not satisfactory or not present and that such training capability, if available, would have a significant payoff. Clearly, devices to train portions of these tasks are in existence (or are currently being procured) so that it was necessary to further clarify the specific events within these HVTs. This was accomplished by the addition of "high value issues and questions" (HVI) to the contract requirements. These address specific "user" questions derived from not only the HVTs but the four primary objective areas. These two techniques, i.e., HVTs and HVIs, provide the study focus and orientation necessary to be responsive to the management team and to Air Force requirements. These tasks were identified and described by members of the STRES Management Team. This team is composed of representatives from each of the major operational, training, R&D and procurement commands, as well as the Army and Navy. It was organized to provide focus and technical advice for this effort on a continuing basis and to enhance the usefulness of the study products.

METHOD

Because of the technical and administrative complexity of the effort, the full STRES program has four major phases. The first phase (Jan 76—Aug 77) was an in-house planning study to identify requirements, assemble the management team, verify study objectives and methods and describe the high value training tasks mentioned above. The phase was also used to prepare and circulate necessary procurement and program materials. Phase II (Aug 77—Dec 79), the current effort, is directed towards historical and on-going operational simulator training and R&D. These data will be analyzed and presented in a series of documents formatted to allow maximum usability. Because it will be impossible to provide information on all the STRES objectives using only historical and operational training cost data, a follow-on (Phase III) of experimental studies is planned to fill gaps in the data base. This Phase will make use of a research plan developed during the Phase II and will take advantage of current R&D findings and operational training requirements to modify the plan as necessary. These experimental studies will be performed using both operational and research simulators. The final phase (IV) will include the total analyses and integration of the data collected during Phases II and III. It will also include recommendations for up-dating of the data base and necessary future efforts.

CURRENT PROGRESS

The Phase II contract was let on 15 August to a three-company team composed of Canyon Research Group as prime, and Seville Research Corporation and United Airlines as subcontractors. Principle members of this team are Chuck Semple and Dennis Sullivan from Canyon, Paul Caro and Wally Prophet from Seville, and Dale Seay of United Airlines. These individuals have a very high degree of collective experience in directly relevant areas.

This Phase (II) has four steps: Step 1, study plan preparation and documentation; review and preparation of literature documentation abstracts. Step 2, on-site data collection, Step 3, analysis and reports, and Step 4, detailed research plan. Development and presentation of the overall study plan was the first task of the first Step. This plan was completed, reviewed and approved by the management team early in the program. Because of the large volume of reports and other types of documentation available relating to the concerns of this study, the contractor was directed to collect, evaluate, analyze, abstract and report those documents that had direct relevance to the study objectives. The purpose of this was to organize, present and define a specific body of data available for use during the study, for use by the members of DOD community as a stand-alone reference, and to ensure that duplication of previously accomplished studies does not occur. The identification, collection and abstracting of documentation has begun. To date, over sixty reports from various sources have been abstracted and hundreds of others reviewed for suitability. The unique feature of these abstracts, and what distinguishes them from the type normally found in reports, is a special section that contains the reviewers comments regarding the presentation format, method, statistical analysis, conclusions and/or applicability of the article in question. The objective is to allow the reader of the abstract access to the major pieces of information in the report without the requirement to consult the primary document. This process will have two major products. First, the technical data contained in the reviewed documentation will be applied, as appropriate, to the analysis and conclusions drawn from the data collected on site. Second, the abstracts will be documented in a bibliography of source materials for those interested in issues related to the objectives of the STRES program. Draft copies of these first abstracts have been distributed to members of the management team for review and comment.

Step 2, On Site Data Collection. This step has just begun. The purpose of this step is to collect and integrate training simulator operational and research information and data, through interviews with government and commercial personnel experienced and knowledgeable regarding objectives of the STRES Program. Recognizing that the published literature is insufficient to adequately describe and characterize the present state-of-the-art in simulator training application, design, and utilization, these data must be obtained from representative organizations presently involved in relevant activities. It is because such broadly stated objectives would be extremely difficult to achieve that the "high value tasks" previously mentioned have additional value. They serve as a focus for the study, and help ensure that primary attention is given to these high payoff training tasks.

Sept 3, Analysis and Reports. This step involves the analysis,interpretation and documentation of the information obtained from the literature review abstracting process and the on-site data collection visits. The goal is to develop models, principles, guidelines and data in a format that can be used by operational training personnel, managers, research and development personnel, and those responsible for support of effective and efficient training programs that use air crew training devices. Although work on this step will not begin for a few months, an in-house mini-study has begun to provide specific user requirement/data format guidance to the contractor team.

Step 4, Detailed Research Plan. Analysis of the data acquired or developed during the previous steps will undoubtedly identify areas in which the present data base is less than perfect. The objective of this step is to prepare a detailed research plan which could provide the kinds of information necessary to resolve these gaps. This plan will not address all areas of simulation training, but will be shaped by the STRES objectives, and high priority needs identified by the operational community. Preparation of the plan will include identification of suitable research facilities, identification of specific information deficiencies, development of candidate research efforts, prioritization of recommended projects and formal plan preparation.

The current phase of the STRES program is expected to be complete in November 1979, with Phase III (experimental studies) scheduled to begin in October of that year.

CONCLUSION

This program, as described, is attempting to bring together many diverse efforts to provide valid and practical guidance for those concerned with aircrew training devices. While it is too early to provide conclusions, it is clear that study objectives are correctly oriented. The operational training community has great need of useful guidance regarding maximizing the value of their current devices and allowing tradeoffs concerning future capability. STRES will obviously not solve all these problems, but with adequate follow-on and updating efforts, has the potential to significantly improve the current data base and add to the understanding of these critical issues.

VISUAL PROCESSING: THE EFFECTS OF PERIPHERAL HEADING CUES IN FLIGHT

Mark F. Godfrey, Bruce A. Smith and Jock C.H. Schwank
Department of Behavioral Sciences and Leadership
USAF Academy, CO 80840

ABSTRACT

This report is a summary of two experiments involving 48 male pilot trainees, that investigated the effectiveness of using peripheral light displays to indicate heading during flight in a GAT-1 simulator. Both experiments showed no significant difference in heading errors between the normal round dial display and the peripheral lights, and both showed that when using the peripheral lights the pilots made more attempts to correct the heading. In the second experiment, a secondary task (digit cancelling), did not change the performance relationship across the displays. These results are consistent with the theory of separate processing for foveal and peripheral vision.

Introduction

Due to increased complexity in modern aircraft, cockpit instrument panels are becoming overcrowded with normal, foveally placed instruments. Pilots have reached their limit in the number of foveal placed instruments they can monitor effectively. Therefore, the Air Force is trying to apply psychological and human factors knowledge to the improvement of flight control monitoring devices.

In response to this, the Aerospace Medical Research Laboratory, through the Department of Behavioral Sciences and Leadership, USAF Academy, is conducting several experiments to determine the feasibility of modifying and then mounting certain flight instrument displays in the extreme peripheral region of a pilot's vision. The effectiveness of this type of display would support the theory that extreme peripheral vision is a separate sense modality from foveal vision (Vallerie, 1967), with direct connections from the peripheral retina, through the vestibular system, to the midbrain.

Recent studies also show that extreme peripheral vision, since it is monitored at the non-conscious level by the mid-brain could be resistant to stress (Leibowitz, 1973), whereas foveal vision, controlled by the cerebral cortex, tends to have tunnelling effects under stress.

Consequently, it may be possible for a pilot to engage in foveal processing during a complex flight maneuver while simultaneously and effectively attending to peripherally presented aircraft information in the form of apparent motion. If peripheral vision is effective in detecting motion cues, one would expect pilots involved in relatively stressful situations to produce lower error scores with peripheral signal viewing than with normal instruments.

This paper dicusses two studies on peripheral processing. Experiment I compared pilots performance under three heading indicator conditions; using the normal round dial, using peripheral lights, and using both indicators concurrently. Experiment II replicated Experiment I, and added an additional workload stress under each condition. Both experiments compared the ability of pilots to maintain aircraft heading over the three display conditions while flying a relatively complex maneuver.

Method

Subjects. Each experiment used 24 different USAF Academy male seniors with normal or corrected to normal vision. All were graduates of the USAF Academy T-41 flight indoctrination program, which is a basic introductory flying program where the student solos and learns several basic maneuvers in 18-26 hours of flying. None held a private pilot license.

Apparatus. In both experiments the pilots flew a Singer-Link GAT-1 Simulator, which closely simulates the T-41 aircraft. The simulator motion capabilities are; 360° of yaw, 90° of bank, and a range of pitch sufficient to attain a maximum vertical velocity of plus or minus 2000 ft min (10.16 m/sec).

For peripheral heading cues, the pilots wore a faceplate, which held two banks of light emitting diodes (LEDs), which had the capability of displaying either a steady light or a yoked in series strobe light. Each LED bank, one on each side of the face, contained five LEDs, 2.11 mm in diameter, mounted in a straight line, and 5 mm apart. In Experiment I, the LED banks were mounted vertically and in Experiment II, they were mounted horizontally. In both experiments the middle LED was along the horizontal meridian and 55° from straight ahead when the pilot fixated on a point 3 meters in front. Total visual angle subtended between the first and fifth LEDs was 14.5 degrees. Presentation of both visual signals was unilateral, left or right, whenever heading was out of tolerance one degree or more. In order to terminate the LED display and thus null an error signal, the pilot was required to make a heading correction toward the side on which the visual stimulation was presented. In the steady mode, the lights were either on or off depending on the heading. In the strobe mode the LEDs flashed faster the greater the heading deviation.

In Experiment II a digit-cancelling device, located on the instrument panel, provided an additional workload stress to the pilots. The device had a single number display that would randomly display a number between 0 and 9, which the pilot was required to cancel using a keyboard of buttons randomly numbered between 0 and 9. After the pilot cancelled a number, a new digit immediately replaced it.

Training. In both experiments the pilots first had a training session where they became familiar with the trainer and proficient to criterion in flying a continuous vertical S-alpha maneuver. This is a relatively complex and demanding maneuver which consisted of alternately ascending and descending 250ft (76.2 m) above and below a baseline altitude of 2000 ft (609.6 m), while maintaining a constant heading of 270° airspeed of 80 mph (35.75 m/sec), and vertical velocity of 500 ft/sec (2.54 m/sec).

The pilots were considered proficient in the maneuver when they could maintain heading within plus or minus 10° of 270° while flying the maneuver under each of three conditions; *normal* (using only compass dial as heading indicator), *redundant* (using both compass dial and peripheral lights to indicate heading), and *peripheral* (using only peripheral lights to indicate heading). In Experiment II, they were also required to be proficient in flying the maneuver with the additional workload stress. Before beginning the experiment, the subjects were instructed that flying was the primary task and digit cancelling was secondary.

Experiment. The experimental session normally occured two days after the training session. The pilots were given several minutes of warm-up before testing. In Experiment I the testing included three four-minute sessions, each under one of the three instrument display conditions previously mentioned. In Experiment II the pilots were run under the same three display conditions, but were also run with and without stress under each condition, making six four-minute sessions. The order of the conditions in each experiment was counterbalanced across the pilots but remained the same for each pilot during the training and experimental sessions.

Dependent variables were obtained using a strip chart recorder. These variables included heading, airspeed, veritcal velocity, and in Experiment II, the number of correct responses to the digit cancellor. Because the strip charts continually recorded the variables,. a further measure of performance was available. This measure, "overcorrections", was based on crossovers from one side to the other of the desired heading.

Results and Discussion

A perfectly performed vertical S-alpha lasts exactly 2 minutes. To allow for two complete maneuvers, each pilot was run under each condition for 4 minutes, with the middle 200 seconds being used for test data. For heading and airspeed, absolute deviations from the desired were sampled at 5 second intervals. The ability of the pilots to maintain a constant vertical velocity was measured from a point at which the standard rate was initially established during climb or descent, until the reverse direction was initiated. Absolute deviations from the standard were also sampled at 5 second intervals for this variable. In all three variables the absolute deviations were averaged for each instrument condition. The number of heading overcorrections were also tabulated along with the number of correct responses to the digit canceller in Experiment II.

To analyze the data a 2TS (Type Signal) X 3TD (Type Display) ANOVA was used for Experiment I and a 2TS X 3TD X 2TST (Type Stress) ANOVA for Experiment II, with the subjects evenly grouped by type of peripheral display. In both experiments, variations in the type of heading display configuration yielded no significant decrement in control of aircraft airspeed or vertical velocity. In addition, there was a non-significant difference between type of lights (strobe vs steady) however, errors in heading were smaller in the steady mode than in the strobe mode.

Figure 1 compares the average absolute heading deviations across the three indicator conditions for Experiments I and II without stress and Experiment II with stress. The only significant difference was between the normal and redundant conditions of Experiment I ($F=3.77$, df – –, *4, $p < .05$), with the redundant condition having a smaller average error than the normal condition. Figure 2 compares the number of overcorrections across the same conditions as Figure 1. In both experiments the average number of overcorrections were significantly greater in the peripheral condition than the normal condition (Exp I: $F=22.13$, df=2 44, $p < .001$ -- Exp II: F-12.46, df=2 44, $p < .001$). This means that the pilots were probably more aware of the errors when using the peripheral lights since they made more overt attempts to correct for the errors.

In both comparisons (heading and overcorrections) there was no significant detrimental effect due to the secondary task of the digit cancellor; however, as seen in Figure 2 there were fewer overcorrections under the stress of the task in all three conditions. These results tend to show that the addition of the workload task did not occupy enough of the pilots attention to cause a significant decrement in performance. Although there was a difference in the number of correct responses to the secondary task between the display conditions, none were significant.

From the results it is apparent that peripheral displays are at least as effective and possibly more effective at indicating heading errors in the cockpit than the normal round dial displays. As stated above, under the peripheral light condition, the pilots were probably more aware of the errors as shown by the increase in overcorrections. Even with the increase in attempts to correct the errors, overall slying performance was not significantly degraded. This indicates that although the pilots were more aware of the error when using peripheral lights, because they did not have to fixate on them, it did not interfere with the normal foveal processing. This tends to support the hypothesis that peripheral vision is processed separately from foveal vision.

Although this study only investigated using peripheral lights to indicate heading errors, they could be used to indicate other instruments in the cockpit. This could be a valuable aid to a pilot because he could monitor several different instruments at once without fixating on them, allowing more attention to a primary flying task. These peripheral indicators could indicate to the pilot when he was out of tolerance on a less critical dimension, thereby prompting him to correct the error.

Leibowitz, H. W. Detection of peripheral stimuli under psychological and physiological stress. *Visual Search*, Symposium, National Academy of Sciences, Washington, D.C., 1973, 64-76.

Vallerie, Leroy L. *Peripheral Vision Displays.* Dunlap and Associates Inc., Darson, Conn., for Electronics Research Center, National Aeronautics and Space Administration, June 1967.

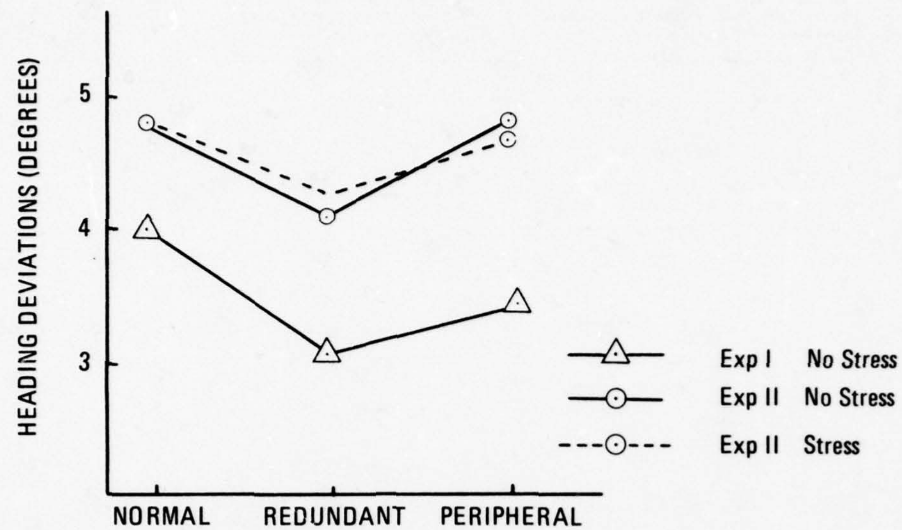


FIGURE 1
Mean heading error deviations across three heading indicator conditions

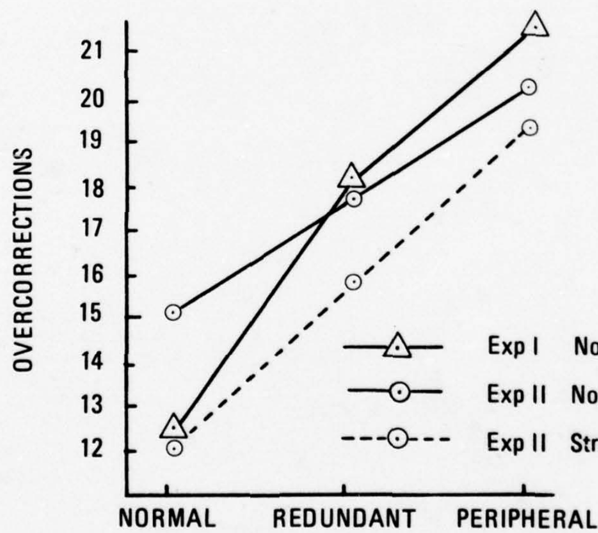


FIGURE 2
Mean overcorrections for heading across three heading indicator conditions

CONSTRUCTING STIMULI WITH KNOWN TRUE SCORES FOR DETERMINING VALIDITY OF RATING SCALES

George W. Lawton
U. S. Army Research Institute
Alexandria, Virginia

Walter C. Borman
Personnel Decisions, Inc.
Minneapolis, Minnesota

Under contract to the Army Research Institute, Personnel Decisions, Inc. of Minneapolis, Minnesota, has developed videotaped samples of job performance. These job performance samples have known "true scores" on several dimensions of the job depicted, and are being used in a continuing program of research on performance ratings. In this research, we are examining the validity of performance ratings as a function of individual differences in raters, rater training, and rating scale format. I want to describe the procedure for constructing these job performance samples in the event that others may find it useful.

There were two phases of the Project I am going to report: *construction* of sample job performances; and *validation* of sample job performance "true scores" against expert judgment.

In outline, the construction phase was as follows: first, behaviorally anchored scales for the dimensions of each job were developed; second, these scales were used to construct realistic job performance profiles; third, scripts depicting different levels of performance were written; fourth, experts familiar with each job evaluated the scripted performances.

In developing the behavioral scales for each job, incidents of job behavior exemplifying different levels of performance effectiveness were content analyzed to uncover the dimensions underlying job performance. Behavioral examples were assigned to the most suitable job dimensions, and examples selected to cover the full range of performance effectiveness for each dimension. The behavioral retranslation procedure of Smith and Kendall (1963) was used to assess reliability of job dimensions and behavioral scaling.

To construct realistic job performance profiles, a matrix of intercorrelations among job dimensions was constructed. This intercorrelation matrix was used, with a Monte Carlo procedure, to produce several job performance profiles.

These job performance profiles were used in conjunction with the behavioral scales to write scripts showing appropriately effective or ineffective performance.

Finally, judges familiar with the jobs depicted, rated each scripted performance. Where there were discrepancies between the judges' opinion and the script, the scripts were modified. The final scripts were videotaped using actors.

With this outline of the construction procedure in mind, let me return to the beginning and fill in some of the details. Two jobs were used in our research: In one set of job performance samples, ratees were depicted as college recruiters representing companies interested in hiring students nearing graduation; in the second set of job samples, the ratee was depicted as a manager holding a problem solving interview with a troublesome employee. These jobs were selected because a relatively small job sample could show most aspects of performance, and because they were very familiar to PDI staff. I will describe construction of job samples depicting different levels of performance for the manager's job. Procedures were similar for both.

Behavioral examples were collected which showed effective, average, and ineffective performance, for a total of 196 examples. Content analysis of these behavioral examples resulted in seven dimensions underlying the manager's job:

- structuring and controlling interview
- establishing and maintaining rapport
- reacting to stress
- obtaining information
- resolving conflict
- developing employee
- motivating employee

Behavioral scaling and retranslation were accomplished together in three steps: first, seven psychologists familiar with the manager's job were given the 196 performance examples in random order and asked to sort them into the seven job dimensions. If an example was assigned to the same dimension by five of the seven judges, it was retained, otherwise it was rejected from subsequent analysis. Second, *from the reliably sorted examples for each job dimension, seven examples were selected which best covered the full range of performance* on the dimension. An example of "very ineffective" performance was assigned a value of "1", "average" was "4", and "very effective" was "7". Third, judges who were not involved previously, were given the scaled examples in random order for each job dimension, and asked to rank order them within their dimension. Interjudge agreement, and agreement with intended rank order were both substantial. Coefficients of concordance, measuring interjudge agreement, were between .94 and .98 for the manager's job (between .86 and 1.00 for the recruiter job). The end result of this phase of development was a behaviorally anchored rating scale for each job dimension.

Next, five expert judges, familiar with both the concept of correlation and the contents of the job, were given the dimensions of the job, and asked to estimate the correlation between each pair of job dimensions on a seven point scale; "7" indicated perfect positive correlation, "1" perfect negative correlation, and "4" indicated zero correlation. Intraclass correlations (Haggard, 1958) which compared variability in different judges' ratings of the same dimension pairs with total variability, were computed at .81 for manager job and .82 for the recruiter job (both $p < 01$). This indicates considerable reliability in this judgment, so mean ratings were computed and converted directly to correlations. Dimensions means of 4 and standard deviations of 1.5 were arbitrarily selected, and, along with the matrix of intercorrelations, used as the basis for Monte Carlo simulation which produced 8 realistic job score profiles for the manager's job, and 8 for the recruiter's job.

These intended true score profiles were used with the scaled behavioral examples, to write 16 scripts, 8 depicting the manager's job, and 8 depicting the recruiter's job. A group of judges used the behaviorally anchored rating scales to rate the performance depicted in

each script. If there were reliable discrepancies between ratings and intended scores for a given script, the script was rewritten to reflect appropriate change. Actors were hired and each script videotaped. This procedure was designed to produce videotapes in which each job performance sample depicted a profile of true scores on each job dimension.

In the next phase of the research, the intended "true scores" of the job samples were "cross-validated" against a criterion of expert judgment. Fourteen industrial organizational and counseling psychologists rated each sample of job performance. Interrater agreement indices, ANOVA measures of convergent and discriminant validity (Kavanaugh, McKinney, and Wolins, 1971), and correlations of expert ratings and intended "true scores", were computed. Prior to viewing the videotapes, each judge was given copies of all scripts for careful study, and urged to become thoroughly familiar with them. Judges were then shown all videotapes and asked to rate each performance on each job dimension.

Intraclass correlations were computed for each job dimension. These ranged from .91 to .98, with a median of .97. These values reflect high interjudge reliability.

The ANOVA procedure suggested by Kavanaugh et. al. (1971) showed that desirable sources of variance, *ratee* (convergent validity) and *ratee X dimension* (discriminant validity), were clearly significant. So also were the *rater* and *rater X ratee* sources, measuring leniency-severity, and halo. Fortunately, the desirable sources accounted for appreciably more variance. A comparison of intraclass correlations from this and other studies showed that these expert ratings had relatively high convergent *and* discriminant validity.

Mean expert ratings were correlated with intended true scores, with results ranging from .42 to .97, with a median of .91. Where there were discrepancies between expert ratings and intended true scores, the high interrater agreement among the experts argued in favor of regarding this as a flaw in the videotaped performance. In these cases, expert ratings were substituted as "true score" values.

The result of this complex development and validation was a set of 16 videotaped job samples with known "true score" values. In research already concluded, these samples have been used to study individual differences variables correlated with rating accuracy. Ongoing research is using them to study rater training, and rating scale format effects.

REFERENCES

Borman, W. C., Hough, L. M., and Dunnette, M. D. *Performance Ratings: an investigation of reliability, accuracy, and relationships between individual differences and rater error*. Alexandria, VA: U.S. Army Research Institute, in press.

Haggard, E. A. *Intraclass correlation and analysis of variance*. New York: Dryden, 1958.

Kavanaugh, M. M., McKinney, A. C. and Wolins, L. Issues in managerial performance: Multitrait-multimethod analyses of ratings. *Psychological Bulletin*, 1971, 75, 34-49.

Naylor, J. C. and Wherry, R. J. The use of simulated stimuli and the "JAN" technique to capture and cluster the policies of raters. *Educational and Psychological Measurement*, 1965, *25*, 969-986.

Smith, P. C. and Kendall, L. M. Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology*, 1963, 47, 149-155.

TABLE 1
"Validity" Correlations Between Experts' Ratings and Intended True Scores

Dimension	Graduate Students Mean Ratings (N=8)	Professional Mean Ratings (N-8)	All Experts' Mean Ratings N=8)	Median Validity for Individual Expert Raters	Range of Validity Co-efficients for Individual Expert Raters
RECRUITER JOB					
A. Creating A Favorable Image for the Company	.89	.94	.93	.84	.68 - .98
B. Organizing the Interview	.82	.74	.78	.71	.41 - .86
C. Providing Relevant Information About the Company	.92	.90	.92	.83	.45 - .90
D. Asking Relevant Questions	.57	.71	.67	.57	.09 - .94
E. Answering Recruitees' Questions	.96	.96	.97	.87	.74 -.96
F. Establishing Rapport with Interviewees	.97	.88	.94	.79	.61 - .98
MANAGER JOB					
A. Structuring and Controlling the Interview	.71	.76	.76	.74	.40 - .86
B. Establishing and Maintaining Rapport	.87	.92	.92	.82	.58 - .93
C. Reacting to Stress	.41	.39	.42	.30	-.31 - .65
D. Obtaining Information	.66	.65	.66	.59	.36 - .83
E. Resolving Conflict	.63	.70	.70	.52	.43 - .91
F. Developing Whipker	.96	.96	.96	.90	.77 - .97
G. Motivating Whipker	.87	.92	.91	.84	.54 - .96

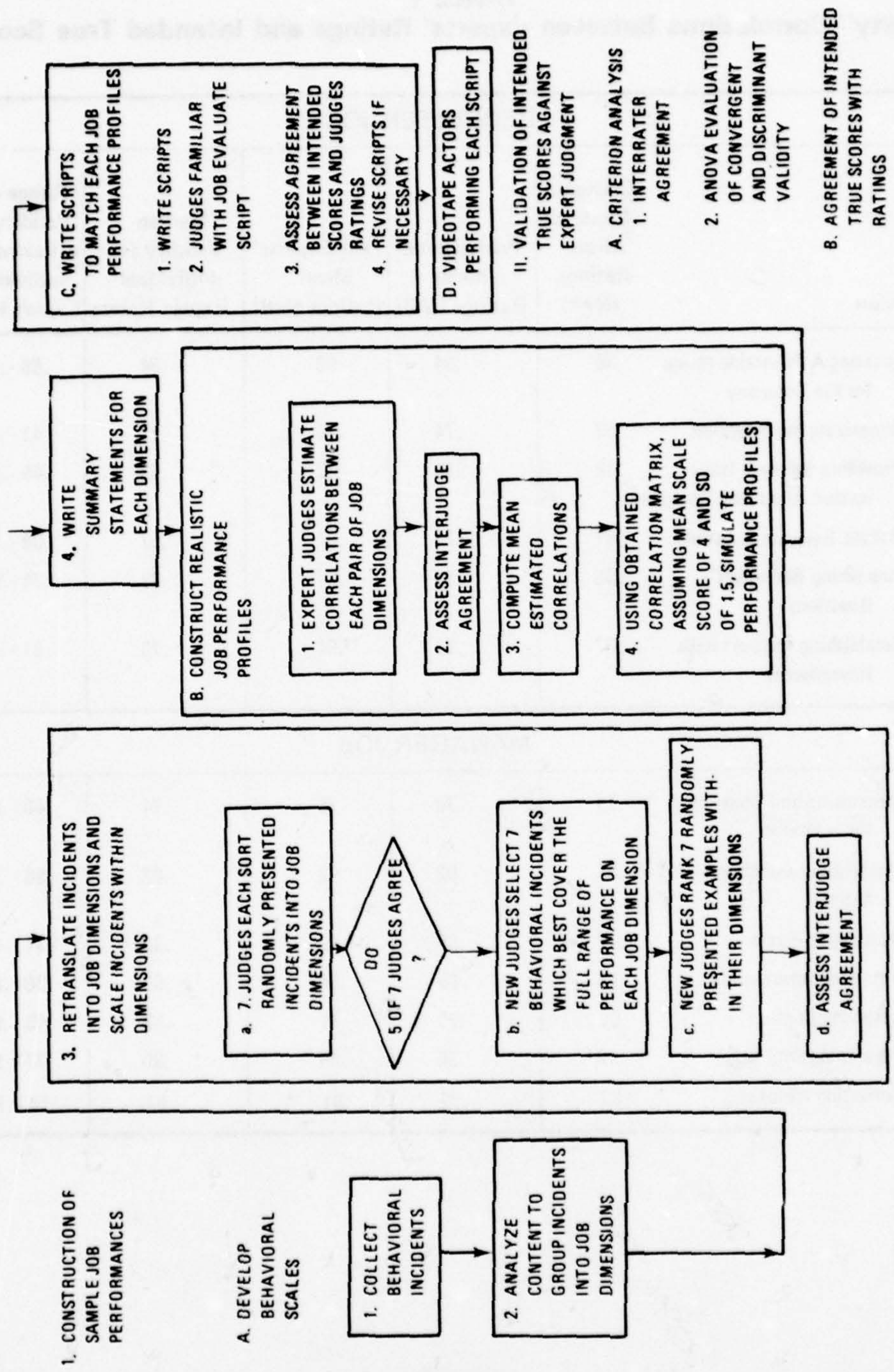


FIGURE 1

General procedure for constructing and validating sample job performances with known "true" scores for use in assessing validity of performance ratings.

PHYSICAL AND PSYCHOLOGICAL STUDIES OF THE CURRENT CHEMICAL WARFARE THREAT: A TRI SERVICE STUDY

James A. Vick
Office of The Surgeon General

ABSTRACT

Intelligence reports indicate that the USSR and the Warsaw Pact countries are extremely well prepared to conduct both offensive and defensive chemical warfare. This real threat has prompted the Army, Navy, and Air Force Surgeon General offices to assess the US/NATO countries position concerning defense against chemical attack, individual and unit training and protective equipment available in case of attack. In addition Medical Operations, Pentagon, has compared the emphasis placed on chemical warfare by the USSR with our current thinking. A two year study has indicated that the US needed to update its protective overgarments, mask, nerve agent antidote, collective protection shelters for handling contaminated patients, detection devices and overall training. An intensive effort by Medical R&D Command, the Academy of Health Sciences and The Surgeon General's Office has yielded excellent results. For example: a new nerve agent antidote has been fielded. STRAF units in the US as well as field soldiers in Europe and Korea now have available an antidote far more effective against G & V agents than those previously available. New protective overgarments and masks are being tested by TRADOC and appear promising. The M51 protective shelter has been successfully tested and is in the hands of the manufacturer. A joint effort between the UK and the US has provided a patient protective wrap to more easily handle contaminated battle field casualties. Furthermore, a new chemical prophylaxis is being tested that would prevent the lethal effects of high dose nerve agent exposure without incapacitating or interfering with the performance of the combat soldier. Extensive psychological and physiological tests have and are being conducted jointly by Army, Air Force and Navy testing facilities to determine the effects of threat, and prophylaxis/therapyregimens on individual and unit performance. These tests include possible effects of prophylaxis and therapy on performance of Army, Navy and Air Force flight crews. Finally the NATO position on chemical warfare is presented along with the current tri service position concerning the overall NBC threat and our defense against it.

Recent intelligence reports indicate that the USSR and the Warsaw Pact countries are extremely well prepared to conduct both offensive and defensive chemical warfare. This threat is both real and psychological. From the real threat standpoint we are aware that there exist numerous Soviet and Chinese installations that are fully prepared to conduct both chemical and biological warfare. These have been shown to be extremely active areas. From the psychological standpoint the sheer knowledge of a real threat puts us in a position in which we must continuously concern ourselves with defending against this type of threat. Even though the United States has ceased all activity of an offensive chemical and biological warfare nature we must still be aware of the threat and be able to defend ourselves adequately in case of attack. Attacks could occur in Korea, or in Germany and could involve the military forces, the US DOD civilians and their dependents. There is also the distinct possibility that an attack could actually occur within the continental United States. In addition, it is thought by some that chemical and biological attack could be linked to a nuclear attack. This could pose a multiple threat with multiple problems; however, most experts are convinced that if chemical warfare is to be used by a potential enemy that it will be employed on a small scale in a densely populated area using a non-perisitent agent. A non-persistent agent is one which produces the desired effects and then is quickly disseminated. Little or no residual contamination remains after a short period of time and the attacking force could easily move in with little fear of resistance and with virtually no damage to the fixed equipment. Chemical and biological warfare are basically designed therefore to be used first as a continuous psychological threat and then also as a real threat to prevent unnecessary destruction yet to accomplish the set purpose; that of killing or disabling the enemy. We therefore must be prepared to defend against it. During the past three years, extensive studies have been conducted in an effort to bring the US defensive posture against chemical and biological agents to a more effective level. This attempt to regain or restore our posture is both in the form of individual protective and collective protection. Individual protection begins in the form of protection over garments; a cloth material impregnated so that chemical agents cannot penetrate. This overgarment is designed to be used in a highly contaminated area to prevent the agent from reaching the skin and causing poisoning through the percutaneous route. The over garment used in conjunction with the protective mask provides optimum protection for troops in the field. Most troops are capable of functioning for from six to eight hours with the overgarment on and the protective mask in place. Psychological and physiological studies have indicated, however, that there are performance decrements associated with their use. The heat load on the individual obviously depends on the ambient temperature, yet treadmill tests have shown that mask plus overgarment can result in troop fall out within 90 minutes at 100 degrees. Treadmill tests have shown that mask plus overgarment in 100 degree F temperatures can result in heat stroke or exhaustion within one and one half hours in troops that are actively performing combat duties. In addition to the mask and the overgarment, the US is conducting studies concerned with a better and more effective antidote to be used in case of chemical or nerve agent attack. Prior to 1973 the only available antidote was Atropine, which was poor at best. Atropine proved to be completely ineffective against those agents which are felt to be the principal threat of the Sino Soviet block countries. As a result of this knowledge and the inadequacy associated with the use of Atropine in treating nerve agents exposure, a new antidote was developed and fielded by the U.S. This antidote has the trade name Combopen. It has been nicknamed TAB which designates the three components of COMBOPEN; Atropine, TMB-4 and Benactyzine. Recent testing in experimental animals indicates that TAB is far superior to Atropine in the treatment of nerve agent exposure. The antidote is given at time of first symptoms such as would be experienced by field troops, exposed to nerve agents and is injected by means of an auto-injection device. If symptoms persist, or worsen, a second COMBOPEN is used. Further treatment, if required, will be given at designated medical installations in which additional antidote and/or oxime could be given. A second concept has surfaced for the protection of troops against chemical attack. This is the use of a prophylactic drug routine. Prophylaxis is a concept whereby drugs are taken in advance of an anticipated chemical attack. This of course requires advance warning, and is not applicable to front line troops, however, with a good advance warning system the 30 or 60 minutes required for effective prophylaxis to work could be attained to protect troops in the rear. If advance warning is indeed available, two chemical routines or regimens are available, both of which have been proven to be extremely successful in preventing either incapacitation or death following chemical attack. The adequacy of the prophylactics has been shown by data wherein animals exposed to lethal dosages of agents were able to survive provided they had been protected by the drug. An interesting psychological aspect to this prophylaxis is that there does not appear to be any significiant performance decrement associated with its use over long periods of time. Humans have actually been administered protective prophylactic doses which are known to be effective against nerve agents for as long as thirty days. They have shown neither side effects nor any

problems in conducting their everyday activities. This concept has received tremendous attention, yet its use must be weighed carefully against problems such as the required early warning. In addition to individual protection, the US and its NATO allies are considering the use of the M51 protective shelter. The M51 would be used in conjunction with field medical installations and is an inflatable shelter with a series of filters. It provides the patient or individual once decontaminated effective protection. The development of the M51 is a joint effort between the Navy, the Army, and the Air Force. The three services are interested in the M51 shelter because of their mutual need for collective protection. In addition the three services are studying all aspects of defense against chemical warfare with the thought of standardizing those medical items which can be of value to all services. There is little question that pilots, whether they be Army, Navy, Air Force, or Marine Corps all share the same concern over performance decrement as a result of mask, protective clothing, prophylaxis or even therapy associated with protecting against nerve agents. Thus, we feel strongly that greater efforts must be made by all services to better defend against chemical attack with the ultimate goal being that of sharing these efforts with our NATO allies.

THE PRECISE MATCHING OF VARIOUS APTITUDES WITH JOBS IN ELECTRONIC-RELATED FIELDS

Thomas J. O'Connor
USAF Human Resources Laboratory

A paper presented at the Sixth Symposium on Psychology in the DoD

In support of the Air Force mission, the Air Training Command (ATC) trains thousands of people. Some of the most expensive training is in the field of electronics. The Air Force has about 63,000 personnel involved directly in the electronic fields and the cost of their training runs into the hundreds of millions of dollars. The Air Training Command estimates that Air Force electronic training costs about a half million dollars per day.

Until now, there has not been any reliable method of identifying precisely the amount of electronic theory needed to perform various jobs. As a result, the Air Force tended to overtrain its personnel. The main purpose of this project is to develop a universal model to evaluate usage of basic electronic principles training.

The criterion used by the model to evaluate electronic theory training is a determination of the usefulness of the training vis-á-vis the performance of assigned tasks in the various electronic career fields.

It is important to be able to identify in a very specific manner those portions of the electronics training which contribute most and least to the performance of tasks. The identification of the relative merit of various portions of electronics training has important implications for managers in the areas of training, personnel classification and testing.

General Background

The model developed in this project is called the Electronic Principles Job Inventory (EPI). The EPI is different from the usual task oriented survey in two major respects. First, the EPI asks two general questions: "What do you do?" and "What electronic knowledge do you use in performing your job?" The usual task survey concentrates on only one question: "What do you do?" The second difference is that the EPI can be administered to anyone who works with electronics. That is, it is general in nature, unlike the usual job inventory which is aimed at a single specialty within a career field.

The EPI is similar to the usual occupational survey in that the data can be analyzed using the Comprehensive Occupational Data Analysis Programs (CODAP).

The EPI contains two sections. Section one has the usual background information, such as rank, command, job title, active military service time, etc. Section two contains the electronic type questions.

A typical job description of an electronic specialty may include a general terminology which uses such verbs as monitors, analyzes, identifies, installs, maintains, troubleshoots, repairs, modifies, aligns, inspects, calibrates, isolates, etc. Two questions may be asked. First, does the action verb mean the same thing for different specialties? Second, even it the action verb does mean the same thing in a general sense for different specialties, does each specialty use electronic knowledge on the same depth and breadth dimension? The EPI results are independent of how one would answer the above questions. That is, the EPI asks questions at such a basic level that the data will yield the kind of information one can use to make decisions without being affected by semantic and communication problems that have plagued previous attempts to deal with the complexity of electronic training.

Brief Description of the Model

In general, electronics courses start with the simple and continually build toward the complex. That is, basic components such as resistors, capacitors, etc. are introduced first and eventually they are combined to form such items as power supplies or motors. The following example illustrates how the EPI determines the utilization of electronic principles training.

The example is for motors only but a comparable section has been written for resistors, capacitors, diodes, transistors, etc. The section on motors starts by asking: Does your job involve any tasks dealing with either alternating current or direct current motors? If the individual answers no, he is routed to the next section of the EPI. If the response is yes, the EPI seeks to determine specifically what tasks the individual performs on motors. The format is a follows:

Do you perform any of the following tasks on motors?

1. inspect . . . . ______YES ______NO
2. troubleshoot down to component parts . . . . ______YES ______NO
3. troubleshoot as far as checking wire connections but do not troubleshoot down to component parts . . ______YES ______NO
4. clean or lubricate . . . . ______YES ______NO
5. operate . . . . ______YES ______NO
6. remove or replace *complete motors* . . . . ______YES ______NO
7. remove or replace motor parts . . . . ______YES ______NO
8. other (specify) . . . . ______YES ______NO

____

From an examination of the pattern of responses to the above tasks, some obvious assumptions about the level of electronic knowledge needed to do different tasks can be made. That is, an individual who removes or replaces complete motors generally doesn't need to know as much about motors as an individual who troubleshoots down to component parts.

The EPI also seeks to determine if the individual performs tasks on specific motor parts such as the field coil, armature, rotor, brushes, slip rings, commutator or pole pieces. Subsequent questions ask whether the individual is involved in determining magnitudes and directions of torque and induces voltages in motors. The final set of questions on motors asks about the types of motors the individual works on (synchronous motor, induction motor, etc.).

The complete data on motors and all the other sections are analyzed to determine field utilization of the electronic principles training.

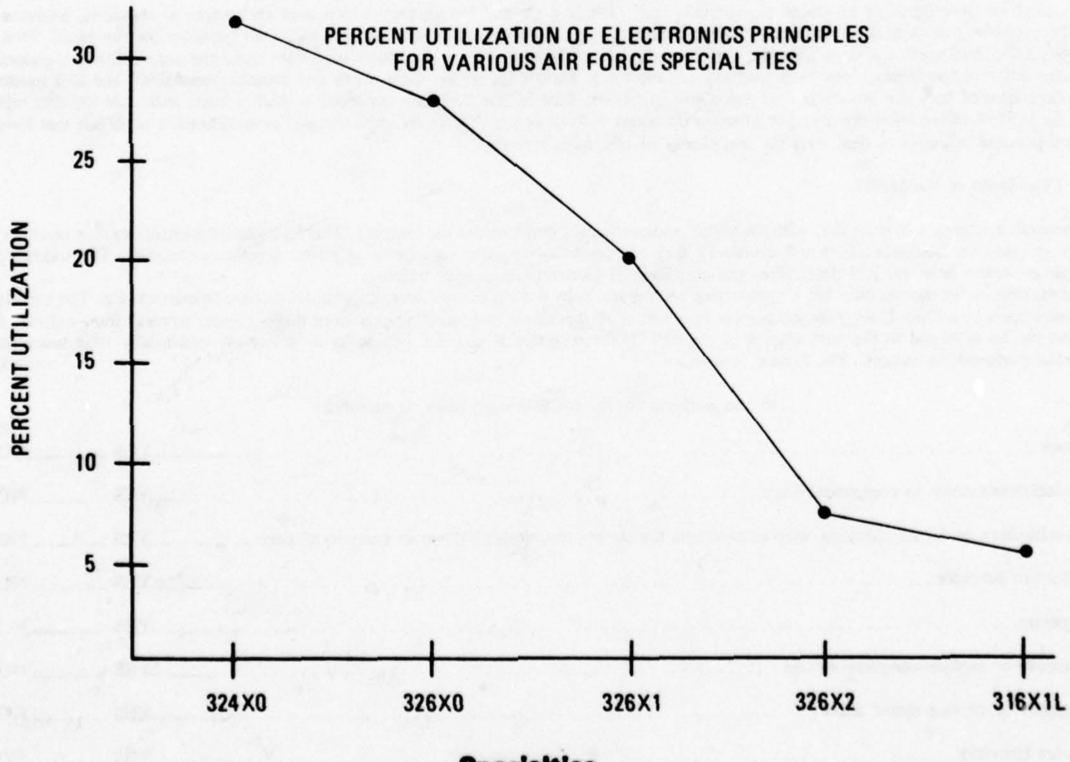
Administration

The Electronics Inventory (EPI) was administered by mail to 1,097 airmen world wide assigned to all shreds of the three 326XX career ladders. This total represents 31 percent of the airmen assigned to these career ladders.

Personnel in the 326XX career ladders are subdivided as follows: 326XOA, Manually Operated Avionics Aerospace Ground Equipment (AGE); 326XOB, Automatic Avionics AGE; 326XOC,F RF-4 Peculiar AGE; 326XOD, A-7D Avionics AGE; 326X1C, Manual Avionics AGE Test Station Operator; 326X1D, Automatic Avionics AGE Test Station Operator; 326X1E, Avionics AGE Operator of Internal and External Penetration Aids; 326X2A, Internal Bomb Navigation, Fire Weapons Control, Digital Computers, and Multi-sensor Displays; 326X2B, Flight Control, Flight Data Recorders, and Integrated Mechanical Instrument Duties; 326X2C, Communications, Navigation, and ECM Systems; and 32692, Integrated Avionics Superintendent.

Results

Figure 1 presents the overall results for the 326XX career ladders. Data for two other career ladders, 324XO (PMEL) and 316X1L (Missile Systems Maintenance, as also shown on Figure 1 for comparison purposes.



Specialties
Figure 1

There were a total of 1,257 electronics principles questions or items in the survey. 326XO career ladder personnel responded "Yes" to an average of 352 items or to 28 percent of the total number of items. The 28 percent is an average figure for all the shreds of 326XO. Figure 1 also shows that 326X1 personnel (all shreds averaged) responded "Yes" to an average of 255 electronics principles items or to 20 percent of the total number of items, while 326X2 personnel (all survey or to seven percent of the total number of items.

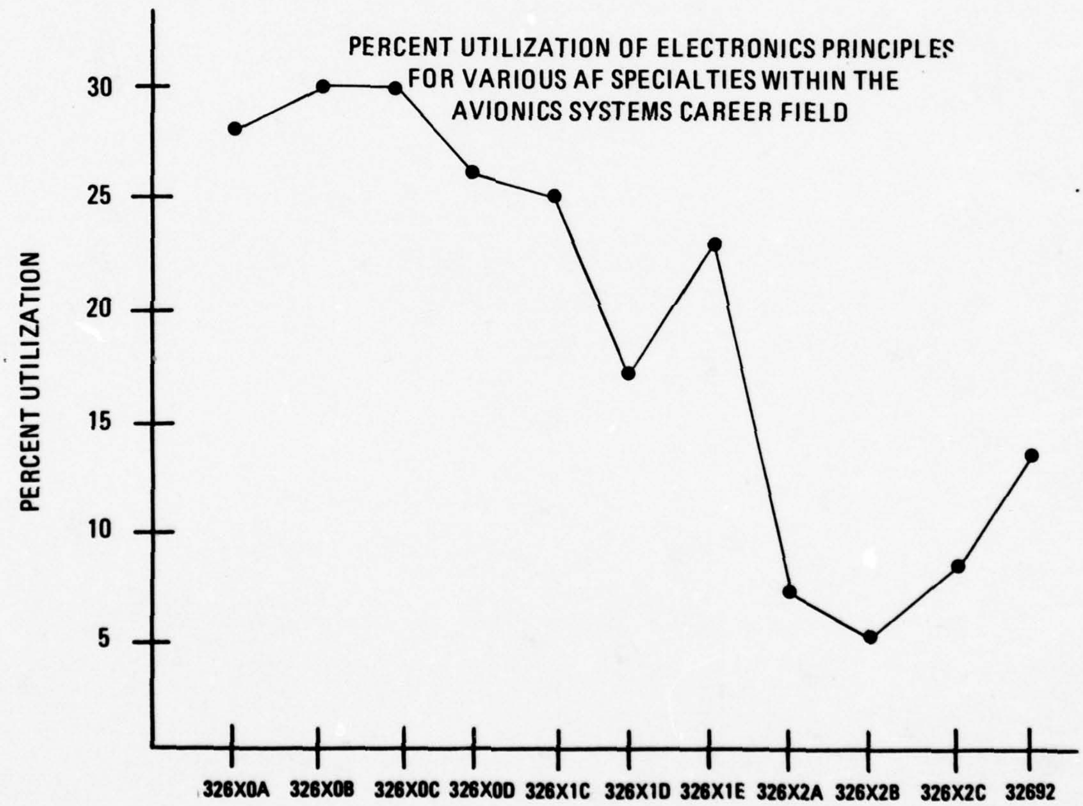
These results, therefore, indicate a wide range of usage of electronic principles among the 326XO, 326X1, and 326X2 career ladders. In addition, Figure 1, shows how the 326XX career ladders compare in field utilization of basic electronics principles with the two other career ladders, 324XO and 316X1L. AFS 324XO personnel responded "Yes" to an average of 401 items or to 32 percent of the total number of items, while 316X1L personnel responded "Yes" to an average of 58 items or to five percent of the total number of items.

Figure 2 shows the percent field utilization of electronics principles for all shreds of 326XX and for 32692 (Integrated Avionics Superintendent). As shown, 326XOC personnel have the highest utilization of electronic principles, while 326X2C personnel show the lowest utilization of electronic principles. It is interesting to note that 32692 personnel show a higher percent utilization of electronics principles than does any shred of 326X2.

Conclusion

At present, the Air Force requires a minimum score of 80 on the electronic aptitude test for the specialists discussed above. There is no concerted effort to place the higher scoring airmen in the more technically demanding jobs. There is also no emphasis on analyzing those jobs which could be performed satisfactorily by airmen who score less than 80 on the electronic aptitude test. The major reason for the deficiencies noted above is that it was impossible to determine, with any degree of accuracy, what the technical demand on each specialty was. The results show that it is now possible to rank order specialties as to the amount of technical knowledge needed to perform each job, and thus, that it is possible to match individuals who receive high electronic aptitude scores with technically demanding jobs, and those with lower scores with less technically demanding jobs.

In the past, the services have had the luxury of selecting from a rather substantial manpower pool. This situation may not go on indefinitely. The Electronic Principles Inventory offers management a unique tool to fine tune its selection process for matching the right man to the right job.



Specialties
Figure 2

MAINTAINING PERSONNEL QUALITY CONTROL AND READINESS THROUGH PERFORMANCE TESTING

Gerald J. Laabs
Navy Personnel Research and Development Center

Personnel readiness, which is an important component of the overall readiness of the U.S. Navy, depends on having enough trained and experienced people to properly operate and maintain the ships, aircraft, and weapons systems in the Fleet. The Navy's personnel system has the responsibility to provide the Fleet with personnel capable of performing their jobs at an adequate level of proficiency. The Fleet has the responsibility to maintain or improve levels of proficiency. However, personnel system decision-makers are often unaware of the source of a personnel readiness problem and cannot take corrective action because typical assessment neither measures job performance nor provides appropriate feedback. Likewise, supervisory personnel are not always aware of performance deficiencies. Furthermore, specific means for correcting performance deficiencies are not usually readily available. This paper reports: (1) a completed project that provides the solution to a personnel readiness problem if the source is the deterioration of skills, and (2) a new project that will provide the solution if the source is the failure to initially acquire skills.

Testing/Training System

Our completed advanced development subproject was based on a very simple premise: if we can discover critical deficiencies of personnel on the job and provide some way of overcoming them, we will improve their on-the-job performance, and as a consequence, improve personnel readiness. In the past, our Center has used a variety of performance tests in the evaluation of training programs and the assessment of fleet performance levels. We have also been active in course design using a variety of systems to deliver individualized, self-paced instruction. These two areas were uniquely combined in a Personnel Readiness Training System which diagnosed critical job-skill deficiencies using performance tests and provided remedial shipboard training in the form of individualized, self-instruction (Anderson et al., Note 1; Laabs, et al., Notes 2, 3, & 4; Winchell, Note 5).

The feasibility of the testing training system was assessed by applying model programs to a variety of operator and maintenance tasks, each associated with a specific rating as follows: (1) Missile Technicians maintaining the Missile Test and Readiness Equipment (MTRE, MK-7, MOD-2), (2) Sonar Technicians operating the AN BQR-2OA passive, real-time frequency analyzer, and (3) Boiler Technicians operating and maintaining the 1200 PSI Stream Propulsion Plant. These ratings and jobs were chosen because they are critical to the mission of their ship or submarine, and there was evidence of performance deficiencies. Therefore, there was potential for immediate payoffs in terms of solving current operational problems in addition to long-term payoffs in terms of generating information of how and where our model program might work.

All three applications were evaluated using the same experimental design (Figure 1). We used diagnostic tests to determine performance levels at the time of a pretest and again at the time of a posttest which occurred from three to five months later depending on the rating being tested. What happened to the personnel during the period between the pretest and posttest depended upon the group to which they were assigned. In each of the three applications, the crews from four ships or submarines were assigned to each of three treatment groups: a Control Group, a Diagnostic Feedback Group, and a Training Group. Members of the Control Group were tested and given only very general feedback on their performance. For example, a Sonar Technician might be told that over all sonar tasks, he used the various controls corectly 80% of the time. Personnel in the Diagnostic Feedback Group were told the specific areas in which their performance was weak. A Sonar Technician might be told that he was weak on the use of the range selector and on the use of gain controls. The Training Group was provided with the same type of diagnostic feedback, but in addition, they were provided with a specifically selected set of instructional materials designed to correct individual deficiencies. It was suggested that the time between the pretest and posttest be partially occupied by working the training materials. There was one additional variable in the Boiler Technician application: two of the ships in the training group were asked to have the assigned materials completed onboard ship; the remaining two were asked to have them completed ashore.

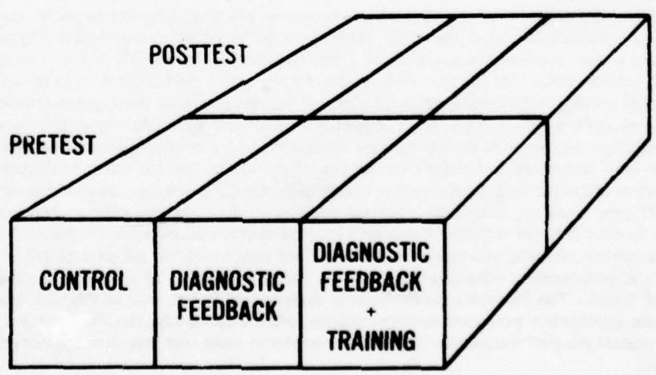


Figure 1. Experimental design for testing training system.

PRECEDING PAGE BLANK-NOT FILMED

For all three applications, the existence of *substantial* performance deficiencies was confirmed by the diagnostic of job-performance tests. The apparent reasons for these deficiencies varied with the area of application. For the Sonar Technicians operating the AN BQR-20, appropriate training apparently had never been given, either in schools or onboard ship, nor had adequate operator manuals been provided. For Missile Technicians maintaining the Missile Test and Readiness Equipment, the nature of the equipment contributed indirectly to the deficiencies. MTRE is a very reliable piece of electronic gear and, during a specific patrol, only one MT is assigned primary responsibility for its maintenance on a rotating basis. Thus, the reliability of the equipment, together with the restricted assignment of maintenance responsibility, provided the setting for skill deterioration . The problems associated with the Boiler Technicians operating and maintaining the 1200 PSI Steam Propulsion Plant have many origins, including insufficient numbers of personnel which impacts on the ship's ability to maintain or improve skills.

A comparison of groups who received feedback-plus-training with those who received feedback only showed that deficiencies were corrected when training was used appropriately. Remedial training materials were fully utilized by the Sonar Technician Training Group which produced significant improvements in job performance. Most members of the Missile Technician Training Group did not complete their assignments. They apparently failed to recognize a need for training because of the excellent reliability of the equipment and the restricted assignment procedures. When the performance of a small subsample of Missile Technicians who had completed training was compared with that of comparable men in the other two groups, the results indicated that had training been completed, performance would have improved. Likewise, we found most members of the Boiler Technician Training Group had not completed their assignments, but an analysis of a small subsample that had completed training indicated that use of training materials would have improved performance. There were evidently two main reasons for nonuse of training materials by the Boiler Technicians: a lack of time to do training onboard ship and the poor shipboard study environment. The Boiler Technicians who were asked to complete training at a shore site completed a far higher percentage of assigned materials.

Comparison of control groups with those groups who received feedback only also showed that feedback alone was not very helpful. In order for feedback to yield improved performance, the individual concerned would have had to act on the information he was given to change his performance. The Diagnostic Feedback Group did show some minimal performance changes as a result of receiving specific information on weaknesses, but even where these were statistically significant, they were not large enough to be of any practical significance.

It should be recognized that feedback in the Personnel Readiness Training System means only providing individuals with information about their inadequacies. The effect of providing such feedback information to others who might make decisions about personnel practices (such as personnel managers or school administrators) was not studied. If the source of a personnel readiness problem is the failure of the personnel system to initially impart the required skills rather than the deterioration of skills, then the testing raining system only provides a short term solution to the problem. In this case, diagnostic testing and shipboard training may improve performance, but cannot remove the underlying cause of a personnel readiness problem.

Quality Control System

If the personnel system is bringing an individual to the job who does not fully meet the requirements of the job, the system components (such as selection, assignment, or training) must be improved to provide a long-term solution to this personnel readiness problem. To bring about the improvements, a procedure is needed to assess the quality of the product being produced by the personnel system and to feed back this quality control information to personnel decision-makers so they can take corrective action. The Performance Proficiency Assessment System, a prototype system we currently have under development, will attempt to provide such a procedure.

In contrast to the Personnel Readiness Training System, the Performance Proficiency Assessment System will be designed to identify deficiencies in the personnel system rather than in individuals. Figure 2 compares the paths to increased personnel readiness through a testing training system and a quality control system. The information needed to initiate quality control actions in the latter system will be obtained from existing sources when available and from the administration of supplemental job-performance tests to relatively small groups of Fleet personnel. The system will be patterned after industrial quality control systems. Of course, the operation of such a quality control system is more difficult than in an industrial setting because the product being produced is not an easily measured piece of hardware. The basic question our advanced development subproject will attempt to answer is: Can a cost-effective system for the measurement of job-performance capabilities be developed, which will provide valid, reliable information on the effectiveness of the personnel processes which bring individuals to their jobs.

Since there are many Fleet measurement efforts currently in progress, a serious question might be raised as to whether or not another effort is really needed. It is important to realize that while many small-scale, partial measurement efforts are ongoing, no *comprehensive* system for the measurement of job performance capabilities of individuals exists in today's Navy. The measurement efforts which yield, or have potential to yield, information about job performance have a variety of shortcomings: (1) Evaluations tend to be made in global terms rather than in terms of specific deficiencies, (2) Typically, evaluations are based upon paper-and-pencil tests or supervisors' ratings rather than actual job-performance tests, (3) Personnel assigned to design and administer these efforts are often not properly trained in performance testing methodology, (4) Some of these programs attempt to test the entire population involved rather than relying on sampling techniques ; consequently, they become overwhelmed by the size of the effort, and (5) Often evaluations are carried out on board ship where various practical and or administrative considerations compromise the types of evaluations which can be conducted. The most serious shortcoming of all these efforts, however, is that the information collected *does not reach* the appropriate personnel decision-makers. Of course, some connections do exist between activities that collect job-performance information and decision makers, but overall, the interconnections are weak. Before precise, reliable information can be obtained concerning the job-performance capabilities of enlisted personnel, all of the deficiencies of the measurement efforts will have to be corrected and the small-scale efforts will have to be integrated into a comprehensive assessment system. The Performance Proficiency Assessment System will accomplish this integration and provide useable feedback information to the appropriate personnel decision-makers. Most importantly, the Performance Proficiency Assessment System will provide for the much needed job-performance testing of Fleet samples to make sure that all of the information required to make personnel decisions will be available.

In summary, our new advanced development subproject will attempt to develop a cost-effective system which will provide for (1) the use of currently collected information whenever possible, (2) the development of needed performance tests, (3) the administration of such tests to appropriate samples of Fleet personnel, (4) the analysis of performance data from these and other available sources and relating of results to specific personnel practices, and (5) the preparation of reports that will provide personnel managers with understandable information concerning the performance capabilities of Fleet personnel.

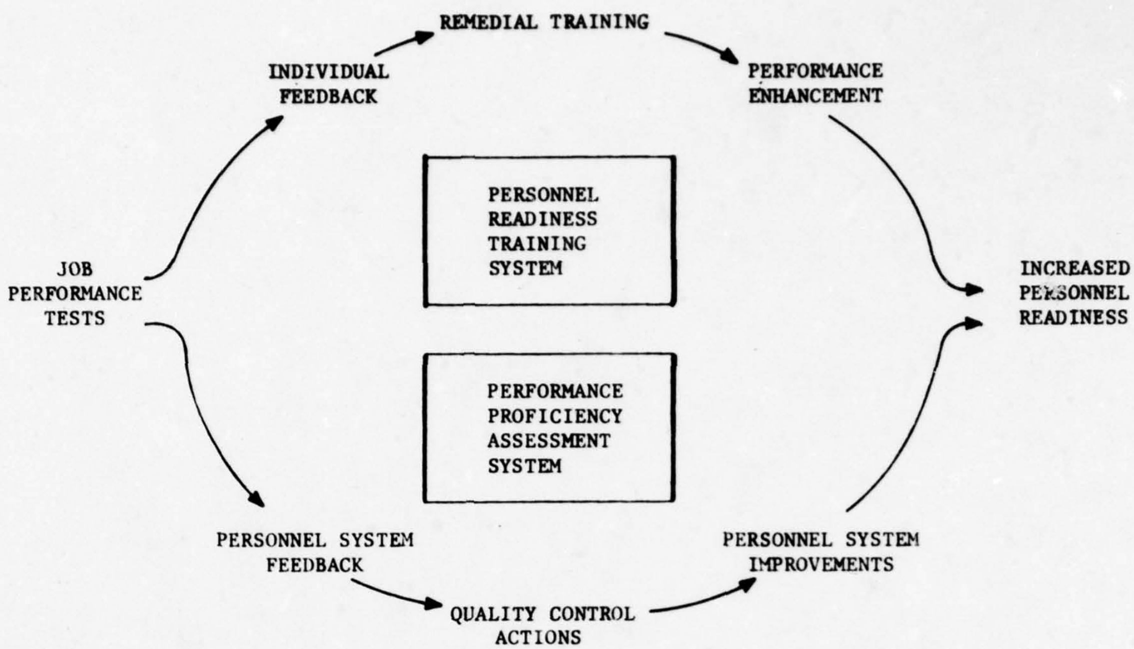


Figure 2. Two paths to increased personnel readiness levels.

The opinions expressed are those of the author and do not necessarily reflect those of the Navy Department.

REFERENCE NOTES

1. Anderson, A.V., Laabs, G.J., Pickering, E.J., & Winchell, J.D. *A personnel readiness training program: Final report* (Technical Report, TR 77-39). San Diego: Navy Personnel Research and Development Center, August 1977.

2. Laabs, G.J., Main, R.E., Abrams, A.J., and & Steinemann, J.H. *A personnel readiness training program: Initial project developments* (Special Report 75-8). San Diego: Navy Personnel Research and Development Center, April 1975.

3. Laabs, G.J., Panell, R.C., & Pickering, E.J. *A personnel readiness training program: Maintenance of the Missile Test and Readiness Equipment (MTRE Mk 7 Mod 2)* (Technical Report, TR 77-19). San Diego: Navy Personnel Research and Development Center, March 1977.

4. Laabs, G.J., Harris, H.T.,Jr., & Pickering, E.J. *A personnel readiness training program: Operation and maintenance of the 1200 PSI Steam Propulsion Plant* (Technical Report, TR 77-36). San Diego: Navy Personnel Research and Development Center, June 1977.

5. Winchell, J.D.,Panell, R.C., & Pickering, E.J. *A personnel readiness training program: Operation of the AN †QR-20A* (Technical Report, TR 77-4). San Diego: Navy Personnel Research and Development Center, November 1976.

AN ANALYSIS OF PERSONAL AND INSTITUTIONAL VARIABLES ASSOCIATED WITH BCT AND 4TH CLASS YEAR USAF ACADEMY ATTRITION

by

Thomas M. Longridge, Jr.
Valentin W. Tirman, Jr.

ABSTRACT

Based on a recommendation by the General Accounting Office (GAO), a longitudinal study of attrition at the USAF Academy was initiated. This report addresses attrition in Basic Cadet Training (BCT), and during the 4th class year. It was concluded that attrition in Basic Cadet Training is largely a function of personal characteristics of the individual, rather than demotivational aspects of training. Reduction in BCT attrition is, therefore, more a matter of better candidate selection than of modifying training.

While personal characteristics continue to play a role in 4th class year attrition, a major difference was evident in the impact of the training environment. The latter was observed to play a significant role in 4th class year attrition, in contrast to the BCT findings. It was observed that an increased effort to establish specific, standardized and publicized goals for the 4th class year training system is needed.

SECTION I
INTRODUCTION

BACKGROUND

In their 5 March 1976 report on Student Attrition at the five federal service academics (FPCK-76-12), the General Accounting Office (GAO) identified a number of factors related to attrition. In particular, their study noted that the percentage of attrition which can be attributed to the academy environment, as opposed to external factors or student characteristics, increases significantly over time, accounting for 1% during Basic Cadet Training (BCT), 11% during the 4th class year and 35% during the 3rd class year.

Their report, which includes a review of the literature on academy attrition, notes that most of the associated research has concentrated on personal characteristics of the attrittee. While the GAO results indicate such factors do play a significiant role in early attrition, their data also indicate that increased attention to the institutional environment is necessary to account for attrition beyond BCT. Accordingly, their study recommended the service academies undertake longitudinal research to study the impact of the academy environment, and to identify the interaction of demotivational aspects with student characteristics relative to attrition. The research reported herein represents the first phase of an effort to comply with that recommendation.

The conceptual model used in this research is similar to that employed by the GAO. Attrition was viewed primarily as a function of two categories of variables and their interaction--namely, individual traits and the academy environment (Walsh, 1973).

As a longitudinal study, the present research report addresses a single group of cadets, the Class of 1980. It is intended that this group will be followed throughout its four year tenure.

The report is divided into two major sections--BCT attrition and attrition during the 4th class year. Because of the space limitations, the present paper will only summarize the methodology employed and the major findings. For a more complete discussion of the latter the reader is referred to Longridge and Tirman (1978).

SECTION II
BCT ATTRITION

METHOD

Consistent with the overall theoretical model, a large pool of potentially attrition related items was generated, based upon group interviews with a heterogeneous cross section of Academy personnel. These initial items were administered under standardized group testing conditions to a computer generated random sample of 276 4th Classmen and 197 3rd Classmen on May 1976. Based on an item analysis of the results, a 150 item Attritional Assessment Instrument, Version I (AAI-I) was created. The latter was administered to the entire Class of 1980 (N=1572) under standardized group testing conditions on the second ay of BCT, and again immediately following completion of BCT. Resignees (N=78) retook the test during their exit interviews, conducted by the Cadet Counseling Center. The AAI-I was supplemented by an additional 71 item instrument (the BCT Supplement), on which basis cadets rated the perceived motivational value of the BCT curriculum.

RESULTS

On the AAI-I,the the most discriminating items between BCT persisters and attrittees dimensions, as follows

Factor 1--Motivation
Factor 2--Attitude
Factor 3--Interpersonal Relationships
Factor 4--Personal Charateristics

When the results from the start of BCT administration were combined in a stepwise multiple regression analysis, with attrition as the dependent variable, a multiple R of .34 ($p < .01$), accounting for 12% of the attritional variance, was obtained. The BCT curriculum received overwhelmingly positive ratings by both BCT persisters and attritees. (Mean persister rating = 4.9, s=.65; mean attritee rating = 4.0, s=.69, along a six point scale from 1 = strongly demotivating to 6 = strongly motivating.)

DISCUSSION

Consistent differences on personality variables were noted for BCT persisters versus attritees and these differences were evident even at the start of BCT. The BCT attritee tends to be an individual who does not feel comfortable in group kinds of activities, but rather prefers to work by himself. Such an individual therefore would not be expected to display high motivation for group types of goals, or goals that necessitate group activity. Nor would his attitude towards those goals, and towards other group members, be positive. As a result, he would be expected to encounter some difficulty in terms of his interpersonal relationships with group members. Such an individual might be expected to be more sensitive to criticism for failure to meet group standards, and he would be less willing to follow orders aimed at the achievement of group goals. Significantly, the BCT attritee tends to rate himself as less competitive than does the persister.

The results do not indicate that substantial modification in the BCT curriculum is appropriate, since BCT attrition appears to be largely a function of personal characteristics *rather than the institutional environment*. Reduction in BCT attrition is therefore more a matter of candidate selection than of modifying training.

SECTION III
4TH CLASS YEAR ATTRITION

METHOD

For the purpose of assessment during the 4th class year, a revision to the 150 item questionnaire was accomplished. Based on the original item analysis, items pertaining specifically to the 4th class year were substituted for those relating specifically to BCT. The revised questionnair (AAI-II) was then administered to cadets of the Class of 1980 who resigned for other than honor violations Class (N = 1271) in May 1977. Academic, conduct and medical eliminees were not included.

RESULTS

The most discriminating items between 4th class year persisters and attritees were observed to fall in twelve categories. These categories are as follows:

Factor 1 - Orientation towards Military Form, Drill and Procedure

Factor 2 - Perception of the Air Officer Commanding (AOC)

Factor 3 - Assessment of the Upperclass

Factor 4 - Career Commitment

Factor 5 - Overall Acceptance of the 4th Class Year Training System

Factor 6 - Orientation Toward Task Success

Factor 7 - Percieved Organizational Climate

Factor 8 - Perception of Academy Education

Factor 9 - Acceptance of Knowledge Testing

Factor 10 - Perceived Extracurricular Opportunity

Factor 11 - Honor

Factor 12 - Study Time

When these results were combined in a stepwise multiple regression equation, with attrition as the dependent variable, a multiple R of .41 (p. $< .01$), accounting for 17% of the attritional variance was obtained. The results of multiple discriminant analysis correctly classified 84% of the Class of 1980 as 4th Class Year persisters or attritees (Longridge and Tirman, 1978).

DISCUSSION

The results of the AAI-II suggest that a review of the effectiveness of the 4th Class Year training system is warranted. While there exist many areas of obvious continuity between BCT and 4th Class Year attrition, a major difference is evident in the impact of the training environment. The latter was observed to play a significant role in 4th Class Year attrition, in contrast to the BCT findings. It appears that an increased effort to establish specific, standardized and publicized goals for the 4th Class Year training system is needed.

REFERENCES

Longridge, T., & Tirman, V.W. *Personal and Environmental Factors Contributing to BCT and Fourth Class Year Attrition*, USAF Academy TR-78-2, USAF Academy, CO., 1978.

Student Attrition at the Five Federal Service Academies (U.S. General Accounting Office FPCD-76-12). Washington, D. C.: Comptroller General of the United States, March 1976.

Student Attrition at the Five Federal Service Academies - Enclosure A: Technical Report on Survey of Factors Related to Attrition (U.S. General Accounting Office FPCD-76-12A). Washington, D.C.: Comptroller General of the United States, March 1976.

Student Attrition at the Five Federal Service Academies—Enclosure B: Review of Studies on Academy Attrition and Related Issues (U.S. General Accounting Office FPCD-76-12B). Washington,D.C.: Comptroller General of the United States, March 1976.

Student Attrition at the Five Federal Service Academies - Enclosure C: Characteristics of Academy Students (U.S. General Accounting Office FPDC-76-12D). Washington, D.C.: Comptroller General of the United States, March 1976.

Walsh, W. Bruce. *Therories of Person-Environment Interaction: Implications for the College Student.* Iowa City, Iowa: The American College Testing Program, 1973.

HUMAN FACTORS ENGINEERING IN THE AIR FORCE:
PROGRESS, PROBLEMS AND PROGNOSIS

Joseph A. Birt, Lt Colonel, USAF
Aerospace Medical Research Laboratory
Wright-Patterson Air Force Base, Ohio 45433

Paul T. Kemmerling, Lt Colonel, USAF
Aeronautical Systems Division
Wright-Patterson Air Force Base, Ohio 45433

Introduction

In this era of increasing costs and dwindling resources it is perhaps timely to be introspective about the field of human factors engineering (HFE)—who we are, where we came from and where we're going. Although HFE has become an important part of the systems process, it is still plagued with problems of identity, resources, technology adequacy and management support. If we could step into the future and look back upon the problems at hand we might see this period as the field's adolescence. Defining HFE technology as "the concrete body of man-machine oriented information, concepts and methods developed for and applied in the design or modification of man-machine systems", (Birt and Kemmerling, Note 1), we will trace the emerging discipline's progress, view its present status and problems, discuss current trends and forecast its future.

Historical Background

The first organizationally concentrated endeavours for the Air Force date to the end of the second World War. An Army directive to the Air Material Command at Wright Field in May 1945 suggested the laboratory study of psychological aspects of equipment design problems. As a result Dr. Paul Fitts who coined the term "engineering psychology" for such work was selected to head a new Psychology Branch in the Aeromedical Laboratory (Fitts, 1947).

Figure 1 traces Air Force organizational roots. The first ten years "engineering psychology" flourished in the Psychology Branch. This research branch provided application consultation and studies for equipment development offices. By 1958 the branch had expanded from 25 to over 95 professionals and the lab had split it into "human engineering" and "training" research divisions.

In 1960 the lab technology development functions and the systems program technology application functions were organizationally separated. Some 30 professional personnel left the lab to form a cadre for full time human factors engineering support activities to aircraft system programs. About that time human factors engineering functions were developing in other hardware and test organizations. Although the numbers of systems support personnel vacillated over time, only the large group in Aeronautical Systems business has maintained its group strength.

In 1960 technology development and applications split was accompanied by the personnel subsystems (PSS) applications management responsibilities for "human performance" as a system component. The PSS concept, as now revised, is dubbed "Human Factors Engineering Management" by AFR 800-15 and can be regarded as an accomplishment in the applications area (Note 2).

Current Status/Problems

A pervasive adolescent characteristic of contemporary Air Force human factors engineering is its identity crisis. "What is the scope of the field?" "Who are we; who should we be; where should we go?" There is more consensus on content, capabilities and trends than many parochial testimonies reflect; however, terminology continues to confuse the players.

The newly formed AFSC Human Factors Engineering Steering Group has begun to attack the identity problem by sponsoring studies of how the technology should be defined, developed, applied and regulated in the Air Force (Birt and Kemmerling, Note 1). For this activity to be fruitful the authors feel it is necessary to explicitly express the emerging man-machine integration (human factors engineering) technology. Likewise an emerging "human factors engineering" discipline must be explicitly nurtured. Future education, training, and career development for the field must receive deliberate near term attention. These and other challenges facing the field are listed in Table 1. AFR 800-15 can serve as the requiring and organizing framework for improved development application of the technology's tools, concepts and data bases (Note 2). A focused effort to practically implement this regulation will enhance the discipline as well as the technology.

TABLE 1

CURRENT PROBLEMS IMPROVEMENTS FACING THE FIELD

- Identity Crises
- Resources Adequacy and Availability
- Demand, Investment, Return Defined
- Technology Tools, Criteria, Methods
- Interface Concepts
- Technology Gaps
- Applications Coverages

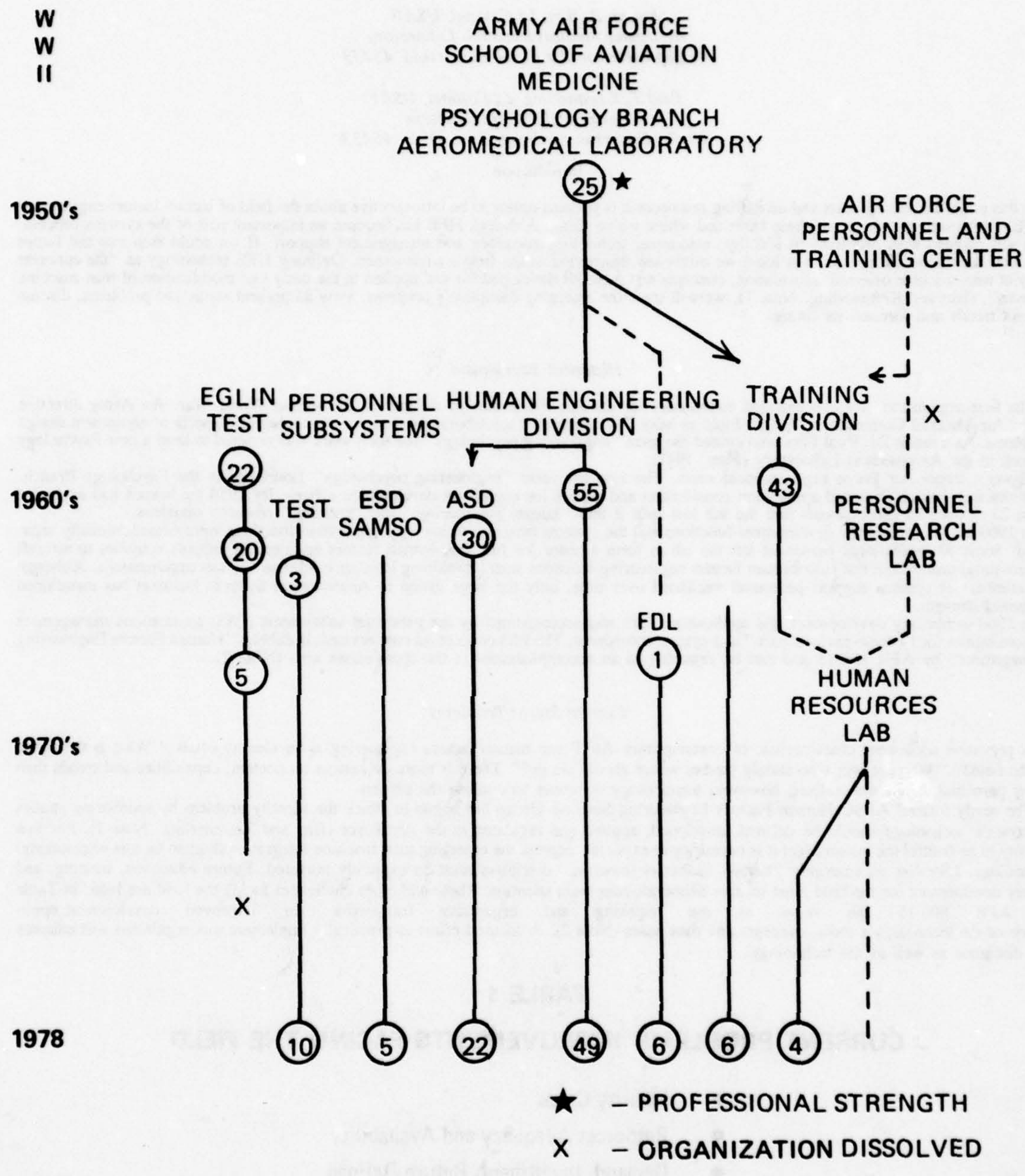


FIGURE 1. ORANIZATIONAL HISTORY OF HUMAN FACTORS ENGINEERING IN THE AIR FORCE

Professional resource adequacy is a major problem. It can best be understood from the trend in technology, applications and management. After the 1960's technology applications split, both the labs and the support groups gained strength until approximately 1966. Subsequently the smaller applications and management groups lost in numbers. Supportive groups at Rome Air Development Center and the Decision Sciences Laboratory at Hanscom AFB had been eliminated by the late 60's. One of the field test's functions at Eglin AFB, Fla. vanished in 1973. Theses trends as well as many of the current problems of the field have been predictable for some time (Birt and Lacy, 1970).

Availability of qualified personnel has troubled the field since the early 1960's. Education cirriculae and training programs always have been loosely defined and controlled for the field. Personnel qualification has largely been a matter of job experience. The large groups with concerned technical management have been able to provide this qualification in depth. The fast expansions of the late 50's and early 60's resulted in opening of Air Force Institute of Technology programs for blue suiters and some graduate opportunities for civilian employees.

The pipelines did not fulfill the PSS demand. In frustration an unfortunate personnel decision was made in 1966 to open a new occupational speciality called Personnel Systems Manager to handle the management aspects of the applications business. No specialized qualifications were required. As a result the positions were filled, some full time and many part-time, by people with little or no training or experience or, in many cases, interest in the field. By this deft move, the requirements for training pipelines vanished as did the education programs and the deliberate input of qualified personnel. Recently that occupational speciality was dropped but the damage done is yet to be repaired.

There is a need to define and institute Air Force education and training programs in this field. The Air Force Academy has begun to take a lead on the education side.

Vacillations in numbers and quality only tell of problems when compared to the demands. Comprehensive information on the demand for the expertise and technology is yet to be compiled. Determining the investment and the return on the investment has been difficult because of the nature of the business. However, it is known that diversity, complexity, cost and responsibilities of the technology and its applications have increased. The development programs in which the technology should be deeply involved are many. Still, major modification programs go without adequate support. Technology development gaps exist.

In the 1976 DOD Topical Review of Human Factors (Note 3), a dollar value of almost 12.5 million was estimated for the technology in the Air Force. Nearly four million dollars were customer reimbursed (Note 3). Since the scope of the field has not yet had uniform definition, accurate figures are difficult to compile. The dollar investments are increasing as are the technology efforts which are being reimbursed by increasing numbers of customers. Requests for support are being turned down for resource reasons by all human factors engineering organizations.

Research and exploratory development emphasis has shifted from "knobs and dials" information to more complex, deeper interface and systems related issues. Yet a need remains for many categories of "knobs and dials" interface criteria for subsystems of varying sophistication. Cognitive information processing implications for interfaces and systems need research attention. Guides and handbooks need to be updated. Improvements and new tools of the trade for each development phase are needed to extend the foundation work of the late 50's and early 60's. Engineering simulation tools need to be explored for improved performance metrics and utility in the RDT&E process.

New interface concepts such as visually coupled systems, voice and eye control, which push related advanced technologies need to be recognized at the conceptual level and mechanisms found to explore such promising concepts, including in-flight evaluations. Routine feedback amongst laboratory technology development, applications, test and operations is needed. Likewise the feed forward process could be improved.

Gaps in our technology for command control communications systems, advanced engineering technology developments, systems modifications, maintainability design, simulation design, safety engineering and operational man-machine problem solution (e.g. accidents) need to be plugged. Most of all, the tehcnology community needs to assure itself that its knowhow, methods, and concepts are appropriately transitioned to the applications world. Resources are critical to these actions.

Applications needs primarily concern the availability of the technology to be applied. Missing data, interface criteria, development methods and specific gaps frustrate and challenge the practitioners. Coverage of programs in the planning and development phases is questionable as is the depth to which our resources will permit coverage on even major programs. The applications efforts are imbalanced with most of the limited resources directed to aircraft versus electronic communication, ballistic space, armament or intelligence systems. Not only are the areas cited as gaps in technology also gaps in application, but there are many technology areas and system programs which go without government and subsequently without industrial effort.

The applications community feels a strong need for quick response study capabilities. Laboratories can to some degree provide this capability. The Aeronautical Systems Division applications cadre has a system simulation facility to address "here and now" questions of alternative interfaces in on-going development programs. The numbers of such questions that the entire human factors engineering community can address must necessarily be limited by the resource capability.

Trends

There are a number of military R&D trends which currently challenge the Air Force human factors engineering community. First, the broad resource picture for military services is not one of overall increases. Reductions in the military component are anticipated throughout the research and development community, yet the human factors engineering area is one which desperately needs the operational experience which blue suit participation can bring.

The diminishing resource picture, of course, has further impact on our technology content. Manpower and maintenance costs weigh heavily on the operational expense of hardware systems. The potential impact of human factors engineering may appreciable affect development costs (Taylor, 1975). Development concerns for life cycle costs should enchance the human factors engineering area. However, here again methodological tools are desperately needed.

We already see trends toward reduced manpower resources influencing our studies in operations and developments. In particular, efforts to reduce the number of crew members on large cargo, tanker and bomber aircraft are percipitating a number of current field and simulation studies.

The Air Force's new prototyping approaches to weapon system technology and development encourage contractor independent design practice. Thus, new specifications which substitute performance for hardware criteria enhance this practice. "Fly before buy" performance

evaluation criteria are needed which necessitate adequate systematic human factors design efforts. We need to develop human factors engineering methods that can be demonstrated to produce cost-effective design practices. We need to learn how to fit man-considerations into these skeleton prototype programs so his roles will not limit the operation of the finally developed system.

The R&D community is tending to move away from full development of new systems. Modifying systems within the constraints of an existing design can be in many ways, a much more difficult human factors engineering task than starting from scratch.

There is an important trend toward wide spread recognition and practice of human factors engineering in many new and unexpected places throughout the development community. With so many groups playing independently in the field there is a tendency to compete in a vast area where cooperation and concentrated development are called for. With the larger quest for data and applications, potential customers are going to diverse sources within government and industry for information and it isn't surprising that diverse information often results.

Management problems are reflected in communication gaps symptoms, the career development training needs and the technology applications needs. Since the mid 1960's there has been a diminishing number of concerned managers at staff levels. Currently there are no human factors engineers as staff members at any major command or Air Force levels. Currently there are no common management (short of major commander levels) or staff functions in the AF which are concerned with the technology and its application. The new AFSC Steering Group is the only forum which currently includes the major known AF systems development players.

Prognosis

We predict that the AF field will outgrow its adolescence and begin to fill a new niche in the research and development world. Through better field definition and scoping, yet-to-be developed education training programs, new collaborative management mechanisms, reemphasized and focused technology development, expanded applications, industrial "imprinting" and improved science-through-operations communications, the field will meet more challenges. It is becoming apparent that a dual qualified engineering and experimental industrial psychology breed will fill the needs of the future. These predictions assume that those of us in the field will unite our views and approaches to solve today's problems and to meet the challenge of the changing R&D environment.

REFERENCE NOTES

1. Birt, J.A. and Kemmerling, P. *Definitions of human factors engineering (HFE) applicable to AF research and development.* Manuscript presented to AF HFE Steering Group September 1977 at Aeronautical Systems Division ENE, Wright-Patterson Air Force Base, Ohio.

2. Air Force Regulation 800-15, 1 October 1974, *Human Factors engineering and management.* Department of the Air Force, Headquarters U.S. Air Force AFPDC, Washington, D.C. 20330.

3. *Air Force man-machine-mission technology (human factors) topical review for DDR&E.* Unpublished presentation material, 21 April 1976. Available from Human Engineering Division, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio.

REFERENCES

Birt, J.A. and Lacy, R.J. Management prognosis in air force human engineering R&D. In Hendricks H.W. (Ed) *Proceedings 1st Annual Symposium, Psychology in The Air Force.* Colorado Springs, Colorado: Department of Life and Behavioral Sciences, United States Air Force Academy, April 1970.

Fitts, P.M. Introduction to psychological research on equipment design. In P.M. Fitts (Ed.), *Psychological research on equipment design.* (Army Air Forces Aviation Psychology Program Reports, Report No. 19). Washington, D.C.: Army Air Forces 1947 (NTIS No. AD-651 795).

Taylor, H.L. Why Human Factors R&D + Systems Development = Improved National Defense. *Commanders Digest,* November 27, 1975, 18 No. 22, 1-8.

A BEHAVIORAL MODEL OF TARGET SEARCH AND DETECTION

Lawrence A. Scanlan and Alice K. Agin
Hughes Aircraft Company

The foundation for a behavioral model of target acquisition performance and the necessary mathematical framework is developed by an integration of diverse research findings and a consideration of the operator's information requirements. The resulting multi-component Markov model is described and a simplified version, used for initial verification of the approach, is presented.

Mathematical models of target acquisition performance are potentially powerful tools for use in the design of electro-optical imaging systems. The extent to which such potential is realized, however, depends upon the accuracy and generalizability of the predictions made by the model. The conditions encountered by operational sensor-display systems are highly complex, and an adequate model must include the characteristics of the sensor, processor, display, atmosphere, target, background scene, and observer. The unusually large number of parameters which the model must ultimately include requires a logical and systematic approach to its development.

Most existing models of target acquisition have used data fitting approaches to obtain equations which predict the probability of detection as a function of time. These models are generally based on a Poisson process and yield an equation with two parameters which can be fitted to observed data. Such models have been successful in predicting performance for abstract targets and for simple or uniform background conditions; however, these models do not generalize to the prediction of performance with realistic, complex backgrounds. This latter case may require equations with three or more parameters. A major problem with the equation fitting approch is the tremendous amount of data which must be obtained. If the final equation is to predict to a large number of conditions, the number of observations required could easily reach into the billions.

An alternative modeling approach which considers the information processing characteristics of the human perceptual system is pursued in the present paper. This more general and more powerful approach is based on an analysis of the operator's task and integrated into a mathematical framework with clearly defined and testable components. The model provides a means for appropriately representing and incorporating both the processing states and the sources of information characteristic of the target search and detection task. The multicomponent model is graphically presented in Figure 1 as a Markov process. The rationale behind this model will be briefly presented below and is more fully explored by Scanlan and Agin (1978).

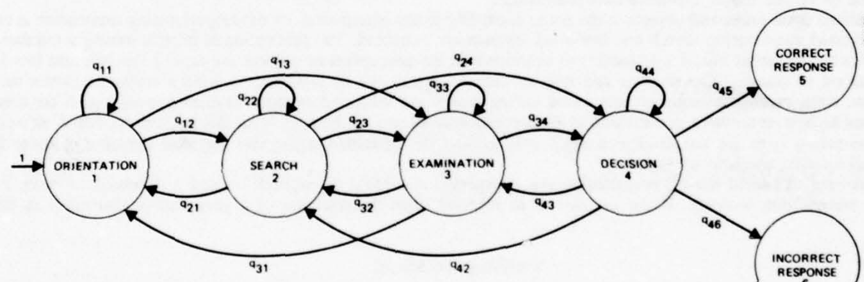


FIGURE 1. MULTI-COMPONENT MARKOV PROCESS MODEL OF TARGET ACQUISITION

Human Information Processing and Target Acquisition

The perceptual informaiton on which operator performance ultimately depends is not a direct function of the data input through the sensory mechnaisms of vision. Rather, it is a complex function of the input data, processing mechanisms, and operator expectation, as indicated schematically in Figure 2. Identification of the relevant information requires consideration of both the task and the processing capabilities of the operator. Further, because the processing required to obtain the necessary perceptual information changes with the expectation of the operator, this internal source of information must also be considered.

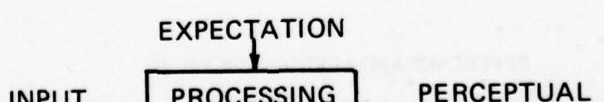


FIGURE 2. ASPECTS OF PERCEPTUAL DATA PROCESSING

Input Data is the visual world as it impinges upon the eye. More precisely it may be defined as the physical, temporal and spatial variation in luminance intensity. In the most general sense, the input data is the total data required to reconstruct a physical, real-world situation and will be a function of a large number of factors.

Evidence from both the neurophysiological and psychological literatures make it clear that the *processing mechanisms* extract features from the input data and that these features are the building blocks from which perceptual information is constructed. Physiological evidence also suggests a hierarchy of feature processing. Low-level features such as size, contrast, color, require little processing time, and are extracted at the more peripheral areas of the visual systems. High-level features which may be composed of several components, such as edges and angles, usually have high spatial frequency components, require foveal processing, and increased time for their extraction. These high-level features are most likely extracted at the visual cortex.

The extraction of features from the input reduces the amount of data dramatically and codes it in a more useful form. However, the feature set requires additional processing to become perceptual information. The features must be selected, ordered, weighted, and combined in a meaninful way. The logical basis for constructing perceptual information from features is a function of *Expectation*. Expectation, the information available to the operator through memory, is a function of the experience and normal perceptual development of the operator, as well as the specific briefing received prior to performing the task.

Perceptual information in the scene. An identification of categories of information within a complex, realistic scene can be obtained by asking observers to describe those characteristics of the scene which might make target detection easy or difficult. A distillation of the responses of 12 subjects yielded four categories of scene information: target, clutter, context, and texture. Each of these can be described in terms of feature characteristics and expectation.

The *Target* will have a set of perceptual features similar to any complex visual pattern. Considerable target acquisition research has been directed toward identifying relevant target features with emphasis on the low-level features of size and contrast; however, under realistic background conditions, these features may not be sufficient for target detection.

Clutter objects are those which are detectable and share some features with the target. In general, a clutter and a target will have similar low-level features but will differ on some high-level feature characteristics. The number of common features between target and clutter, the proximatness of clutter to the target, and the number of clutter objects are factors which can influence detection performance.

Context refers to those terrain objects or areas which have a systematic and meaningful relationship with target location. Normally, context objects seem to facilitate the search process, because they offer the operator information about areas where targets are likely or unlikely to occur, thus, indicating areas the operator should reasonably search or ignore.

Texture refers to areas with relatively uniform or recurrent elements of high spatial frequency and low contrast. The effect of texture on performance is expected to be relatively small.

Scene information and target acquisition. The previous discussion considered the influence of object features and expectation on perceptual information. These aspects of perception can be related to the four states of target search and detection shown in Figure 1. The initial perception of low-level, global scene features and the rapid eye fixations of the operator over a large area of the scene, characterize *Orientation*. The outcome of orientation is a simple global scene description and a general strategy for searching selected areas. Expectation and context are predicted to be the major influences on orientation.

The rapid eye fixations upon areas and objects in the scene according to the search strategy developed during orientation is called *Search*. Objects are briefly fixated upon during search and low-level features are extracted. The perception of objects having a number of low-level features in common with the target elicits a transition to examination; the perception of objects not having the relevant low-level features results in a continuation of search. The number and type of clutter objects can be expected to have a major influence on search.

In the *Examination* state, candidate target objects found during search are subjected to more detailed processing. If the candidate target object has the relevant high-level features, a transition to the decision state occurs. If the high-level features are not found, search is resumed. Fixations in this processing state are confined to a small area around the candidate target and the total duration of these fixations will be longer than fixations characteristic of search.

The high-level features extracted during examination are interpreted according to expectation and a *Decision* to make a response or continue search or examination is made. If the decision is to respond, then the selection of a response is also made in this state.

Verifying the Model

To test the model, it is necessary to estimate the Markov model parameters on the basis of operator behavior as a function of task variables. The locations of eye fixations on the scene is an appropriate method because they are good indicators of the information which is supported by results reported in the literature, and such measurements provide the necessary duration, sequence and probability estimates required to compute the Markov model parameters.

Initial verification of this behavioral approach has used a simplified model because the orientation and decision states are more difficult to identify and because they are expected to require much less time than the search and examination states. The resulting two-component model combines orientation with search and decision with examination. Search is characterized by short duration fixations while examination fixations are longer and include repeated fixations to an area, as well as clusters of sequential fixations in a common area.

The simplified Markov process model has been evaluated under two types of stimulus conditions using eye movement measurements. Abstract scenes provided highly-controlled features for all objects in the scene and eliminated context information. Realistic scenes provided more complex and variable features for all objects, as would occur under natural conditions. The focus of the tests was on the validity of the underlying assumptions about object features, processing components, and expectations. These experiments provide support for the model approach and are detailed by Scanlan and Agin (1978).

REFERENCE AND ACKNOWLEDGEMENT

Scanlan, L.A. and Agin, A.K. A behavioral model of target acquisition in realistic terrain. Culver City, California: Hughes Aircraft Co., Display Systems Department, Technical Report P78-70, February 1978.

The research presented was supported by the Visionics Technical Area of the Night Vision and Electro-Optics Laboratories, United States Army Electronics Command under contract number DAAK70-77-C-0013. Mr. Richard Flaherty monitored the contract.

THE DEVELOPMENT OF HEAD-DOWN DISPLAYS FOR LOW-VISIBILITY LANDING

Richard S. Jensen and Nicholas M. Simonelli
University of Illinois

This paper compares the results of a simulator experiment in which experienced pilots flew landing approaches using various computer generated display configurations with an aircraft experiment in which similarly experienced pilots flew landing approaches using a periscope display. Results of this comparison suggest that properly designed computer generated cockpit displays could be developed to replace contact vision in aircraft landing tasks. The crucial step in such display designs is the development of augmented visual cues presenting directly observable indices of desired and actual performance in a pictorial format.

The design of a head-down display for use as a monitor during automatic landings or as a manual take-over display in the event of automatic system failure requires a compromise between established display magnification principles and instrument panel space limitations. The compromise of display magnification principles may be possible without serious effects on performance through a combination of pictorial display principles, augmented visual guidance cues, predictor symbology, and/or pursuit display configurations.

The principles of pictorial display magnification were established in a series of experiments at the University of Illinois in the late 1940s and early 1950s (Roscoe, Hasler, and Dougherty, 1966). In preliminary experiments these experimenters found that the apparent distance of objects viewed through a periscope on a flat ground glass screen is greater than the perceived distance to the same objects viewed directly. A magnification factor of 1.29 was required to achieve equivalent distance judgments. In a study of aircraft landing performance using this device the authors found that significantly better performance in terms of both mean touchdown error and variability was achieved with a magnification value of 1.20, compared with .86 and 2.00. A tendency to overshoot was observed for magnification factors less than 1.20 and to undershoot for magnification factors greater than 1.20.

This magnification principle presents a real challenge to the designer of head-down (or head-up) displays for aircraft landing. To be of any value in maximum crosswinds the outside visual angle presented on such a screen must be greater than twice the required crab angle for these cross-winds. If not, the scene directly in front of the airplane would not appear on the screen. The required screen size for a 40 degree outside visual presentation viewed from a distance of 22 inches with a magnification of 1.29 would be 28 inches across. A screen of this size inside the cockpit is clearly out of the question. Therefore, a compromise of the display magnification principle is required and other display principles must be applied to correct for the expected performance decrement.

The principle of pictorial realism (Roscoe, 1968) says that the display should present a spatial analog of the real world. The Roscoe, et al. (1966) study demonstrated that pilots could fly the airplane as well with a pictorial display properly magnified as they could seeing the world directly. There is good reason to expect that a properly designed computer generated display in a pictorial format could produce similar results (Eisele, 1978). Furthermore, a computer generated display has the capability of presenting in three dimensions the desired flight path and predicted future positions of the aircraft relative to the desired flight path. If properly designed, this display would permit pure pursuit tracking of the desired flight path by the controlled symbols representing the predicted system states.

The initial study of a research program to develop a three-dimensional display for low visibility landing (Simonelli, 1978) was an experiment to determine the effects on pilot performance of displays presenting the projected flight path (PFP) information for straight approaches from an initial offset. It has been well established, that presentation of such "preview" information improves two-dimensional tracking performance as compared with displays showing present error only (Leonard, 1953; Poulton, 1954). The reason for this improved performance is that preview information permits the operator to anticipate changes in the function being tracked. To effect a comparable improvement in three dimensional tracking requires the design of projected flight path symbology with perceptually convincing distance cues. In this study displays presenting PFP information were compared with runway only, conventional ILS, and ILS plus PFP presentation.

METHOD

Subjects

The subjects in this study were 16 male University of Illinois flight instructors. Participants varied in flight experience from 285 to 1193 total hours. Participation was voluntary, and no monetary payment was either offered or made to any participant.

Apparatus

The data were collected in a Singer-Link GAT-2 configured as a light twin-engine airplane. The GAT was used with its "washout" motion system providing two degrees of movement--roll and pitch. The display, located on the right side, was a 20-cm (diagonal) Conrad video monitor. Viewed from 56 cm, the display subtended a horizontal visual angle of 17 degrees. The outside horizontal visual angle presented on the display was 30 degrees resulting in a magnification factor of 0.57.

Experimental Design

The experimental design consisted of complete factorial crossing of four variables: displays (four types), crosswind velocity (10, 15, 20, and 25 kts), crosswind direction (90 degrees right and left), initial position lateral offset (2000 feet left and right). The display factor was a between-subject variable; four subjects randomly assigned to each display group. All other factors were within-subject variables. Each pilot flew 16 approaches composed of all possible combinations of crosswind velocities, crosswind directions, and initial position offsets. The order of the 16 conditions was counterbalanced among the 16 subjects such that each condition was preceded once and followed once by every other condition.

Four different display configurations were used. Display 1, shown in Figure 1, is the complete projected flight path (PFP), runway (RWY), and ILS scale. Display 2 presented the same format without the ILS. Display 3 presented the runway only (RWY). Display 4 presented the ILS scale only (ILS).

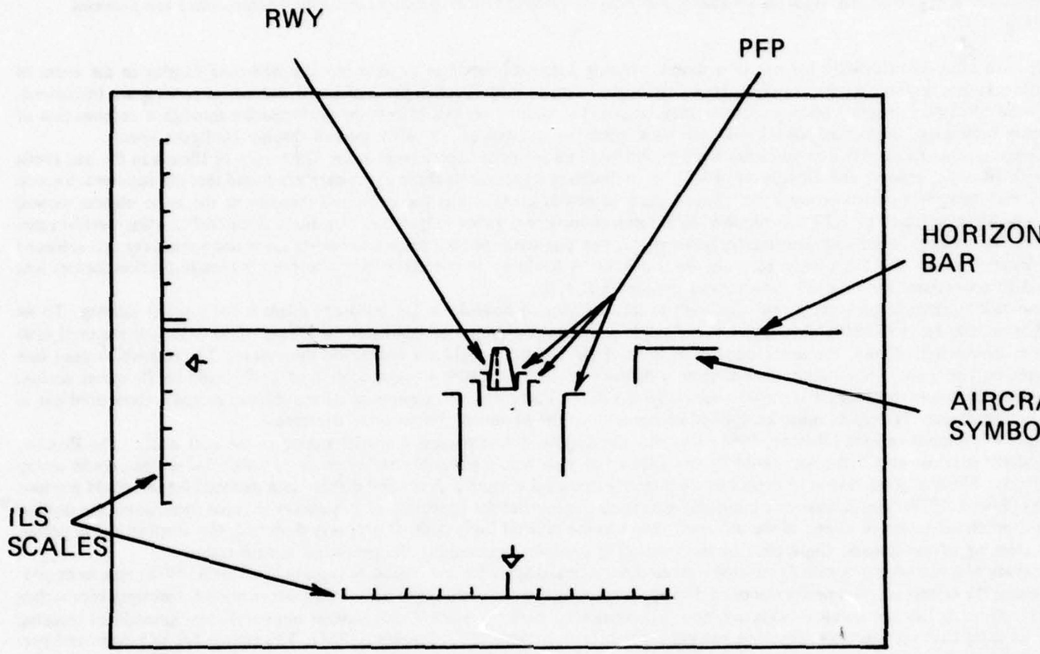


FIGURE 1. DISPLAY 1 SHOWING PROJECTED FLIGHT PATH, RUNWAY, AND ILS SYMBOLOGY.

Procedure

Following a 30-minute familiarization session, each subject flew 16 approaches all in one session. Pilots flew the simulator from the right seat and were asked to intercept the localizer and establish a three-degree glidescope as soon as "comfortably possible." Each approach was started from initial positions 2000 feet either right or left of the localizer on a heading of 360 degrees (the same as the runway). Before each approach, the simulator was set in approach configuration; 130 knots, flaps at 10 degrees, gear down, mixture and propeller controls set for landing. Data were recorded automatically on magnetic tape as each approach was made. Dependent measures were lateral error from the localizer and vertical error from the glidescope.

RESULTS

Of particular importance to the landing task is the delivery window at the runway threshold for each of the displays tested. Figure 2 presents the 1σ lateral and vertical error windows for each of the four displays at 500 feet from the runway threshold. At this point the nominal approach path is 78 feet in altitude, and there remains 1500 feet to fly until touchdown (at 130 knots, about 6.8 secs). One can quite clearly see from this figure that the best vertical and lateral performance was achieved with the ILS and PFP combination. The runway, by itself, yielded both highly variable performance and a tendency to fly above the glideslope; which translates into runway overshooting. This overshoot tendency is partially corrected with the PFP and greatly improved with the vernier ILS scales ($p < .05$). Lateral performance was best for the combination PFP and ILS and worst for the ILS by itself. Thus, the effect of the pictorial display is to reduce lateral error, and the effect of the ILS is to reduce vertical error.

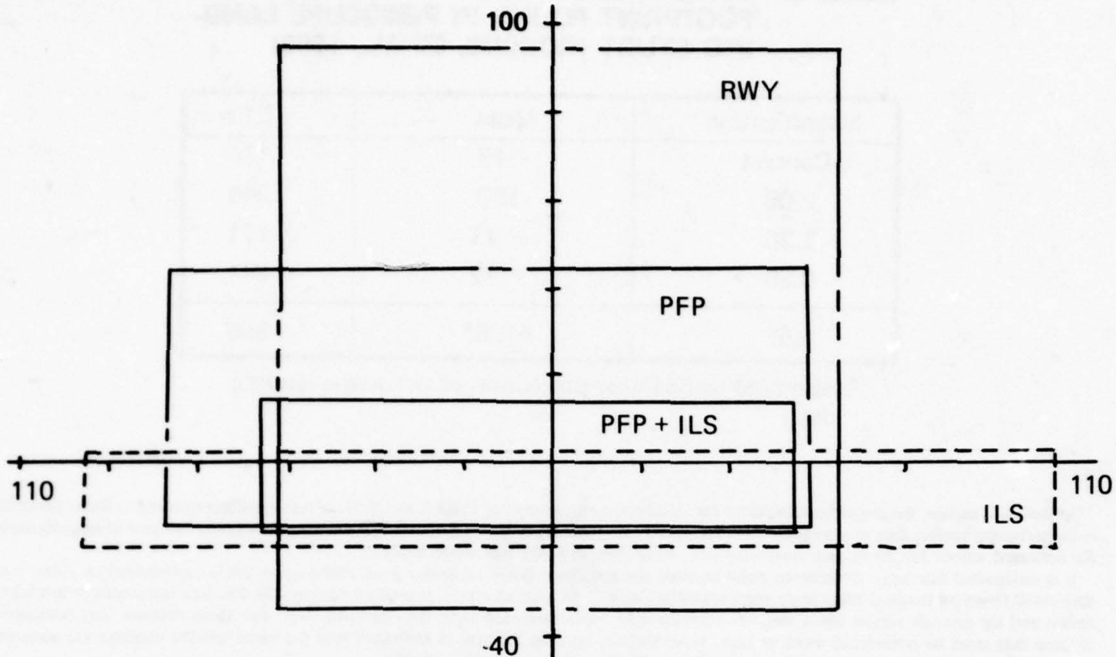


FIGURE 2. 1σ LATERAL AND VERTICAL ERROR WINDOWS FOR FOUR DISPLAYS

DISCUSSION

Although differences were found in pilot performance of the final approach task as a function of the display configuration, the question remains: is preview, ILS scales, or some combination of the two sufficient to correct the display deficiency caused by less than optimum display magnification? A rough comparison of these threshold error window results projected forward to touchdown with the touchdown footprints found in the Roscoe, et al. (1966) study provides some useful information.

Table 1 presents the estimated touchdown footprint in longitudinal and lateral dimensions from these simulator data. For comparison, Table 2 presents the longitudinal dimensions of the touchdown footprint for simulator data was found by taking the average of the worst case forward projection (assuming no improvement from 1500 feet to touchdown) and the best case forward projection (assuming variable error will continue to decrease to touchdown at the rate of decrease found between 3500 feet and 500 feet before the threshold).

TABLE 1. PROJECTED TOUCHDOWN FOOTPRINT IN FEET FROM ERROR STATISTICS AT 1500 FEET TO GO IN THE SIMULATOR WITH COMPUTER GENERATED LANDING DISPLAYS (AVERAGE OF MAXIMUM AND MINIMUM ESTIMATES)

	Longitudinal Axis		Lateral Axis	
	Mean	+1 σ	Mean	±1 σ
PFP + ILS	0	244	0	54
ILS	-166	181	0	97
PFP + RWY	+273	474	0	64
RWY	+563	1062	0	55

TABLE 2. LONGITUDINAL AXIS OF TOUCHDOWN FOOTPRINT FOUND IN PERISCOPE LANDING STUDY (ROSCOE, ET AL., 1966)

Magnification	Mean	$\pm 1\ \sigma$
Contact	- 19	118
2.00	-152	246
1.20	+ 11	171
0.86	+ 72	241
0.57	+128*	305

*estimated using linear projection of periscope landing data.

As indicated earlier, the magnification factor for all displays represented in Table 1 was 0.57. If one is willing to accept a linear projection of the periscope landing data to a magnification factor of 0.57 (Roscoe, et al. suggest only that the mean is a linear function of magnification), the estimated values for mean and variability are +128 feet and 305 feet respectively.

It is recognized that large differences exist between the simulator flown computer generated display study represented in Table 1 and the aircraft flown periscope display study represented in Table 2. As outlined above, to provide comparable data both the simulator touchdown results and the aircraft results for a magnification of 0.57 were estimated from the observed data. For these reasons, any comparison of these data must be considered weak at best. Nevertheless, because the task in each case was the same and the displays are somewhat comparable, the aircraft study provides a plausible validation for the simulator results.

In terms of comparable visual cues for spatial orientation, it is suggested that, of the four computer generated displays used, the runway display (RWY) best represents the aircraft flown periscope display. Although the runway display contains the most useful orientation cues (runway outline and centerline) isolated by Eisele (1978), one would not expect to find the simulator landing performance to be as good as aircraft landing because there may be other useful cues in the actual picture that are not present in the computed line drawing of the runway. Nevertheless, the RWY display will be considered the baseline display around which symbology would be added to achieve parity with contact flight performance.

Considering the rather large differences between the experiments represented in Tables 1 and 2, some very striking similarities are evident. For example, the overshoot tendency predicted by the periscope data for magnification <1.20 are clearly evident in the RWY data and only partially corrected with the addition of the projected flight path. However, the further addition of the ILS scale appears to eliminate any systematic tendency to undershoot or overshoot even with a magnification of 0.57.

The effects of longitudinal variability of the PFP and ILS are equally dramatic. The addition of the projected flight path reduces longitudinal variability to one-fourth of that found for the runway alone. In fact, with the PFP and ILS symbols, longitudinal variability is reduced to levels near those found in the periscope study and considered safe. However, a word of caution concerning the use of the ILS scale is in order in view of the lateral error statistics shown in Table 1. Apparently, the ILS scale by itself fails to provide adequate lateral steering information. These data suggest that pictorial information such as that provided by the RWY or PFP is needed for adequate lateral steering control in the landing situation.

Although the vertical data seem to indicate a requirement for the ILS scale, there are two reasons why an alternative should be sought for the display of this information, especially when the approach path is curvilinear. First, the lateral error data mentioned above show how effective pictorial information when combined with directly observable indices of desired performance (the vertical orientation of the runway centerline) and actual performance (amount of deviation from a vertical orientation of the runway centerline with the wings level). Apparently, the lateral ILS scale could be eliminated with little or no loss in lateral steering performance. If similar indices of desired and actual performance could be found for the vertical dimension within the pictorial format, one could expect to eliminate the need for the vertical ILS scale as well.

Second, although the ILS scale may be located on the same display as the pictorial information, it does not satisfy, perceptually, the definition of display integration for three reasons: 1) it is located ouside the foveal vision area and requires an eye movement to observe it directly, 2) it requires different and more complex mental transformations to establish required control movements from the observed information, and 3) the integration of ILS scale information with pictorial information either requires additional mental transformations, or it requires that the two sources of information be processed separately. All three of these reasons suggest the existence of a less than optimum mental workload, a possibility for serious control error, and a need for a true integration of vertical error information into the pictorial display.

One way to achieve the goal of a pictorial display with directly observable indices of desired and actual performance in both vertical and horizontal dimensions is to use predictor symbols in conjunction with the flight path guidance symbology. Properly designed, the predictor symbol would provide the index of estimated error in both vertical and lateral dimension for direct comparison with the projected flight path, the index of desired performance. This type of symbol arrangement would permit pure pursuit tracking, a goal of all tracking display designs because it eliminates nearly all mental transformations.

This predictor symbology would also provide computer aiding for required control movement anticipations, a feature that has been found to reduce learning time and performance variability to apparently absolute minimums in some situations (Warner, 1969). This combination of predictor symbology with projected flight path information in a pursuit format may be sufficient to offset the performance decrement due to a less than optimum magnification factor, but additional research is needed to determine whether or not this is true.

REFERENCES

Eisele, J.E. The isolation of minimum sets of visual image cues sufficient for spatial orientation during aircraft landing approaches. Unpublished Master's thesis, University of Illinois, 1978.

Leonard, J.A. Advance information in sensorimotor skills. *Quarterly Journal of Experimental Psychology*, 1953, *5*, 141-149.

Poulton, E.C. Eye-hand span in simple serial tasks. *Journal of Experimental Psychology*, 1954, 47, 403-410.

Roscoe, S.N. Airborne displays for flight and navigation. *Human Factors*, 1968, *10*(4), 321-323.

Roscoe, S.N., Hasler, S.G., and Dougherty, D.J. Flight by periscope: Making takeoffs and landings; the influence of image magnification, practice and various conditions of flight. *Human Factors*, 1966, *8*, 13-40.

Simonelli, N.M. An investigation of pictorial and symbolic aircraft displays for landing. Unpublished Master's thesis, University of Illinois, 1978.

Warner, J.D. *A fundamental study of predictive display systems.* NASA Report CR-1274, Washington, D.C., 1969.

HUMAN FACTORS IN TERRAIN MAPS:
RESEARCH FINDINGS AND METHODOLOGICAL DEVELOPMENTS

John P. Farrell and Lawrence M. Potash

Abstract

Map reading is a complex perceptual-cognitive task and in recent years there has been an increasing amount of cooperative research between psychologists and cartographers to facilitate the task of the map reader.

Presenting the three dimensional aspects of land, or terrain relief, on a two dimensional surface to make an accurate and easily understood map is a difficult problem and an optimal solution has not been reached. Contour lines are the most widely used technique to portray terrain relief but many map users have difficulty visualizing terrain with this method. Three experiments on techniques to enhance the perceptual accessibility of terrain information are presented. Results favor supplementing countour lines with larger tints. The development of a research methodology for work in this problem area is also described.

Maps and other topographic displays are a large and increasingly important research area in human factors. The area can be roughly divided into two kinds of problems: 1) those focusing on the map product itself, and 2) those focusing on the user and the processing of the information displayed. A partial listing of the problems relating to the map itself that require psychological research include: appropriate scale, format, and symbology. Symbols may refer to lines, areas, points, and indicate natural or man-made landmarks. Traditional symbols on paper maps may be inappropriate for automated displays but the use of different symbols may produce negative transfer effects. The use of color raises issues regarding hue, brightness and saturation. Another severe problem involves clutter or presentation of extraneous information that detracts from user performance.

This paper focuses on experiments on techniques to improve the perceptual accessibility of terrain information on maps. A secondary purpose of the paper is to discuss some of the methodological problems involved in psychological or human factors research with maps and present the somewhat innovative methodology used in these studies.

The most commonly used method for depicting terrain relief on maps is contour lines. Contour lines are lines drawn through all points of the same elevation on a given landform with the contour interval between the lines reflecting the range of elevation found on the terrain. Contour lines have largely replaced other methods of depicting relief because they provide the most precise quantitative information about terrain relief, however, there are a number of difficulties with contour lines. Some of these difficulties are cartographic: cliffs do not show up well, small terrain features are not represented, and rough terrain with extreme relief such as that found in many parts of the Western United States often requires more contour than map space permits (Marsden, 1955). A more serious deficiency of contour lines is behavioral: many map users have difficulty visualizing terrain when contour lines are the only guide (Skop,1958).

Since it would be undesirable to sacrifice the precision afforded by contour lines, the problem arises as to how contour lines may be supplemented to enhance visualization of the relief portrayed. This problem is a complex one because the addition of material on a map may create clutter and detract from its effectiveness. This problem is a complex one because the addition of material or information to a map may create clutter and actually lead to a decrease in the effectiveness of map product. In the current study two supplementary forms of relief portrayal, shading of slopes, and layer tints were used in conjunction with contour lines to determine the degree to which these techniques enhanced or interfered with the perceptual accessibility of information portrayed by contour lines. Other variables considered are experience level of the map user, and interactions between map format and complexity or ruggedness of the terrain.

A Secondary purpose of the program was to develop a research methodology for work in this problem area. Traditional techniques for the evaluation of new map products consist of surveys of user preference and field studies. Both of these techniques can have shortcomings. Preference surveys are always problematic unless they are carefully designed, and in regard to maps, if the user has not had actual experience using the map, his judgment may lack validity. Previous studies have shown that users often perform better with their less preferred map formats (Wheaton, Zavella, and Van Cott, 1967 and Hill, 1974).

Field studies for the evaluation of maps are also problematic. Some examples of the difficulties in conducting field research on maps are: if only one terrain is used, results may be terrain specific, and seasonal effects must be balanced. These and other variables make field evaluations of maps expensive and difficult to control. As a result of these problems with traditional approaches to map evaluation, an approach, which is described in the method section, was developed where the information extracted from the map product is measured.

Method

Subjects

The subjects consisted of 100 Army officers and NCO's. Previous experience with standard maps employing contour lines ranged from relatively low to high.

Description of Relief Format Assessment Test.

A test was constructed to measure how quickly and how accurately relief information was extracted from a given map format. The problems chosen were selected on the basis of the literature on relief portrayal, consultation with personnel from the Topographic Products Design Branch of the U.S. Army Engineering Topographic Laboratory, and a task analysis. On various map segments, the following problem types were used in the Relief Format Assessment Test:

(1) Identification of landforms such as hills, valley-drawn, spurs, depressions, saddles.
(2) Ridge and valley identification.
(3) Slope identification, i.e., whether the slope is concave or convex, etc.
(4) Determination of the four highest and four lowest grid squares on a 64 grid square map.

(5) Determination of the elevation of various points.
(6) Determination of the vertical profile of various landforms.
(7) Terrain visualization.
(8) Defilade or intervisibility.

The map segments chosen for use in the test were 1:50,000 scale chosen from "off the shelf" maps with the supplementary techniques being added for the appropriate groups. Four combinations of elevation range-contour interval were represented.

Procedure

Experiment I

In experiment I, 48 officers and NCOs with a moderate to quite high experience level were divided into 3 groups tested and compared on the Relief Format Assessment Test. Group A used only contour lines, group F used contour lines supplemented with shading, and group C used contour lines supplemented with layer tints. A two-way analysis of variance for unweighted means and unequal n's was used. Interactions between maps were tested with experience level of user and also with the complexity of the terrain.

Experiment II

In this experiment 30 relatively inexperienced second lieutenants were tested with the Relief Format Assessment test to determine optimum format with a group not contaminated with previous experiment. A two-way analysis of variance for unweighted means and unequal n's was used.

Experiment III

In this experiment 22 officers with moderate of high experience levels were tested on the Relief Format Assessment test and then measured on a field task where they were required to report the coordinates of their own location at several points and also the location of various targets. Field performance scores were correlated with test scores using Spearman rho correlation coefficients. Reliability indices of the eight subtests were obtained and a factor analysis is underway.

Results and Discussion

Experiment I.

On most of the problems slope shading was the poorest of the three map formats. There were fewer significant differences between contour lines alone and contour lines supplemented with gradient or layer tints, although there was significantly less time needed to locate high and low elevations when contour lines were supplemented with layer tints. There were significant interactions between experience level of user and map format. Results in general favor supplementing complexity interactions have not been found.

Experiment II.

Results of experiment II confirmed the effects found in experiment I and indicate that slope shading actually impedes rather than enhances the perception of terrain relief. This finding indicates that this is a general effect on the military map reading problems used in the experiment, and the effect occurs regardless of the experience level of the user.

TABLE 1
CORRELATION OF TEST SCORES WITH ERRORS IN FIELD PERFORMANCE

PROBLEM TYPE	FIELD PERFORMANCE		
	SELF LOCATION	TARGET LOCATION (U)	TARGET LOCATION(S)
LANDFORMS	-.30	-.10	-.34
RIDGE VALLEY	-.14	-.18	-.27
SLOPE	-.32	-.13	-.20
Hi-Lo	-.26	-.49	-.30
SPOT EL.	-.28	-.47	-.22
PROFILE	-.04	-.20	-.33
VISUALIZATION	-.24	-.29	-.63
DEFILADE	-.31	+.01	-.31

Experiment III.

Correlations between the test and the field tasks indicate that the measures used have reasonably good predictive validity.

Summary and Conclusions

Findings on the experiments clearly demonstrate the need for empirical testing of cartographic products. Slope shading, a technique which appears to enhance perception of terrain relief, was found to actually increase the amount of time required to extract terrain information from a military map. It is likely that this effect is due to clutter.

Effects of the three techniques of portraying terrain relief, unenhanced contour lines, contour lines supplemented with shading, and contourlines supplemented with layer tints are presented in detail for eight military map problems.

Since survey of user opinions do not provide a valid assessment of the effectiveness of a map product, and field testing is very costly in terms of time and research dollars, the research methodology used in these experiments should provide a valuable tool for the evaluation of military map products.

REFERENCE NOTES

Hill, A.R. Cartographic performance: An evaluation of orthophotomaps. London: Royal College of Art, Final Technical Report, 1974.

Marsden, L.E. Shortcomings of contour lines, *Surveying and Mapping*, 1955, XV, 2, 182-188.

Skop, J. Investigation of user requirements pertinent to relief portrayal. Washington, D.C.: U.S. Army Map Service, Corps of Engineers, Final Report, 1958.

Wheaton, G.R., Zavala, A., and Van Cott, H.P. Enhanced photo-map evaluation study, Washington, D.C.: American Institute for Research, Technical Report AIR-F-15-4, 67-TR, 1967.

EVALUATION OF THE ORGANIZATIONAL EFFECTIVENESS EFFORT IN THE ARMY: EVALUATION DESIGN

Dan Popov
U.S. Organizational Effectiveness Training Center

ABSTRACT

The U.S. Army's Organizational Effectiveness (OE) Effort began in July of 1975 after a two and one-half year experimental Organizational Development (OD) pilot study. An integral part of the effort is an evaluation of the Army-wide program to provide: information on the program's progress, guidance for the modification of training, information on which to base policy decisions, guidance on field requirements for organizational effectiveness, information as to what the Organizational Effectiveness Effort acutally does, and what it costs to achieve this.

These requirements necessitated a complex evaluation design. This paper outlines the social system modeling procedures developed to meet this need, as well as the methodology used to build detailed descriptive data into a *logic-based*, large scale computer model of the Army's Organizational Effectiveness Effort.

INTRODUCTION

The U.S. Army's Organizational Effecitveness Training Program began in July 1975 after a two and one-half year experimental Army Organizational Development pilot study which was judged worthy, by the Army, of having principles and techniques which if taught appropriately to selected personnel, could enable a positive impact on overall Army operations. Since July of 1975 the OETC has graduated over 300 Organizational Effectiveness Staff Officers (OESOs) who are now assigned at over 70 locations world wide. Some Army installations have been involved with the implementation of OE procedures for almost 3 years. An inherent part of this Army oriented internal OD consultant program is the evaluation of the entire OE effort to provide: information on the progress of OE; guidance for modifying and updating of the training program; information for policy decisions and writing of OE doctrine; and guidelines to D.A. for field requirements for OESOs. In addition, two primary questions had to be addressed, what does the OE effort acutally do, and what does it cost to actually accomplish this.

Due to these rather diverse needs, and the need to produce both final results and continual practical information dictated a combination of an action research program conducted on an ongoing basis, and a basic research program that will produce the answers to the key evaluation questions. The ongoing effort had to be set up to provide timely answers to questions of interest to policy makers, and information to practitioners in the field to enable them to conduct more effective operations. A confounding factor in this entire process is the diversity of the efforts being established world wide. Each program functions autonomously designed to meet the needs of the particular installation to which it is assigned. This rapidily resulted in a situation where the only thing in common from one installation to another was the basic philosophy inculcated during the OESOs training.

PROGRAM EVALUATION DESIGN

The traditional approaches to program evaluation are rarely adequate when approaching a program as complex as the Army's Organizational Effectiveness Effort. In addition, information of use for policy planning is often judged more on the basis of its immediate utility rather than on the scientific rigor of the design that generated it. In order to meet the demands of both utility and accuracy, it was necessary to develop a design that produced results useful in ongoing decision making while maintaining an acceptable form of scientific validity. The concept chosen to meet this need was large scale social systems modeling using a logic based analysis procedure. To handle the breadth and complexity of the data this procedure was designed to facilitate computer based information processing capability.

To begin the design it was necessary to develop a model of the OE effort sufficient to organize both data collection and analysis, and reflect all critical aspects of the acutal functioning of the OE effort. The assessment-focused modeling procedure developed views the Army's OE Effort as an unique combination of the major components: The Organizational Climate of the unit under consideration; The Organizational Effectiveness Staff Officer; and The Organizational Effectiveness Process used by the OESO to assist the unit in achieving a greater level of effectiveness. Each of these components interacts with the other two to produce the outcomes of the effort. These outcomes are controlled by what the organization's climate will allow, what the OESO is prepared to do, and what the OE process predisposes in those circumstances. This interactional model is graphically depicted in Figure 1.

Within each of the components of the interactional model of OE are a series of critical elements or issues, all of which have a bearing on the critical balance of the model. The elements critical to the organizational climate are reflected in Figure 2.

The elements viewed as critical to the OESO's functioning are outlined in Figure 3.

The elements that comprise the organizational process component are outlined in Figure 4.

As the model is interactional in nature, it is broad enough to handle the complexity of an OE effort in a division and yet narrow enough to provide useful information.

PLAN OF DATA ANALSYS

We considered two basic approaches, as outlined by Bennett (1964), to analyze the information gathered during the evaluation effort.

> There are two ways of dealing with complex, or multivariate, problems. One is to introduce arbitrary simplifications so that we can use the techniques of analysis that may be available. This is the mathematical approach. The other is to accept the complexity as an irreducible element in the situation and search for a structure or pattern that will enable us to examine it as a whole. This is the *systematic approach.*

AN INTERACTIONAL MODEL OF THE ORGANIZATIONAL EFFECTIVENESS EFFORT IN A UNIT/ORGANIZATION

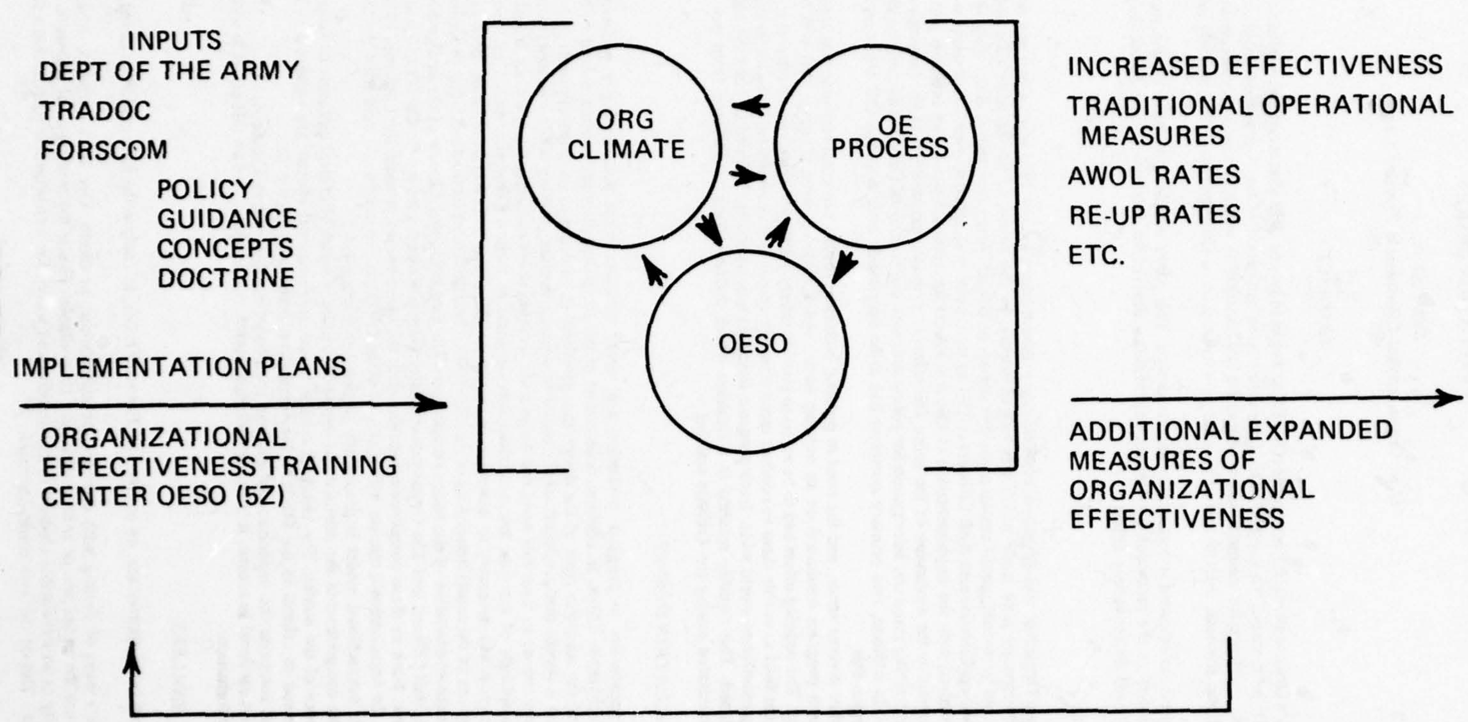


FIGURE 1

ORGANIZATIONAL CLIMATE

Command Support System
External Resources Available
Previous OE History/Understanding of OE
How Things Are Done
Mission
Functional Support System
External Demands/Control of Time
History of/Openness to Change
Personal Support System
Size (ratio)
Access Points
OESO Staff Channel

FIGURE 2

ORGANIZATIONAL EFFECTIVENESS STAFF OFFICER

Credibility

Knowledge

Skills

Experience

Personal Attributes

FIGURE 3

ORGANIZATIONAL EFFECTIVENESS PROCESS

A set of planned behaviors that assist the Commander to maximize the effectiveness of his/her organization to accomplish the assigned mission by increasing the organization's ability to:

Respond to needs for change.
Communicate openly.
Solve problems.
Direct behavior toward a goal.
Balance social exchanges.
Enhance self-esteem
Monitor its own functioning.
Correct non-productive approaches.
Open options and permit choice.
Maximize effective/efficient use of resources.
Enhance individual competence.
Enhance individual commitment.

FIGURE 4

In the process of comparing the two approaches a number of fundamental differences became apparent. The first of these differences rests in the basic underlying questions addressed by each approach. This difference is illustrated in Figure 5.

COMPARISON OF THE TWO MODELS OF ANALYSIS (I)

	TRADITIONAL/ MATHEMATICAL	SYSTEMATIC
BASIC DIFFERENCE IN THE UNDERLYING QUESTIONS I.E.,	IS IT THERE?	WHAT IS THERE
	ARE THE OESOs GETTING GOOD COMMAND SUPPORT?	WHAT COMMAND SUPPORT ARE THE OESOs RECEIVING AND HOW HAS THIS SUPPORT AFFECTED THE OE EFFORT?

FIGURE 5

As the underlying questions addressed determine the *form* of the expected answers, the different approaches lead to different types of expected answers. This comparison if reflected in Figure 6.

Finally, in an action research paradigm, answers, in and of themselves, are not necessarily useful. A comparison of the two basic approaches in terms of the utility of their answers for policy purposes is outlined in Figure 7.

COMPARISON OF THE TWO MODELS OF ANALYSIS (I)

	TRADITIONAL/ MATHEMATICAL	SYSTEMATIC
POTENTIAL ANSWERS	YES, or 60% OF THEM ARE	HERE ARE THE VARIOUS FORMS OF COMMAND SUPPORT THE OESOs ARE RECEIVING AND THEIR INFLUENCE(S) ON THE OE EFFORT.

FIGURE 6

COMPARISON OF THE TWO MODELS OF ANALYSIS

	TRADITIONAL/ MATHEMATICAL	SYSTEMATIC
USE FOR POLICY PURPOSES i.e., IF I WANTED TO ASSIST THE OESO TO ESTABLISH THE OE EFFORT IN MY ORGANIZATION WHAT CAN I DO?	GIVE THE OESO GOOD COMMAND SUPPORT	YOU COULD DO THESE THINGS 1. 2. 3. AND AVOID, IF POSSIBLE 1. 2.

FIGURE 7

Admittedly this analysis is overly simplistic, but points to the reality that for action oriented research of the type we are conducting where the ultimate for evaluating the results of the research is its utility at the time, the systematic approach offered more potential. In addition, there is an added benefit in the process of systematic data processing. In a traditional mathematic approach, data management tends to be reductionistic in nature with at least some of the basic data being lost during analysis. On the other hand, data management in the systematic approach is expansionistic in nature allowing for complete data retrieval at any point. This is graphically illustrated in Figure 8.

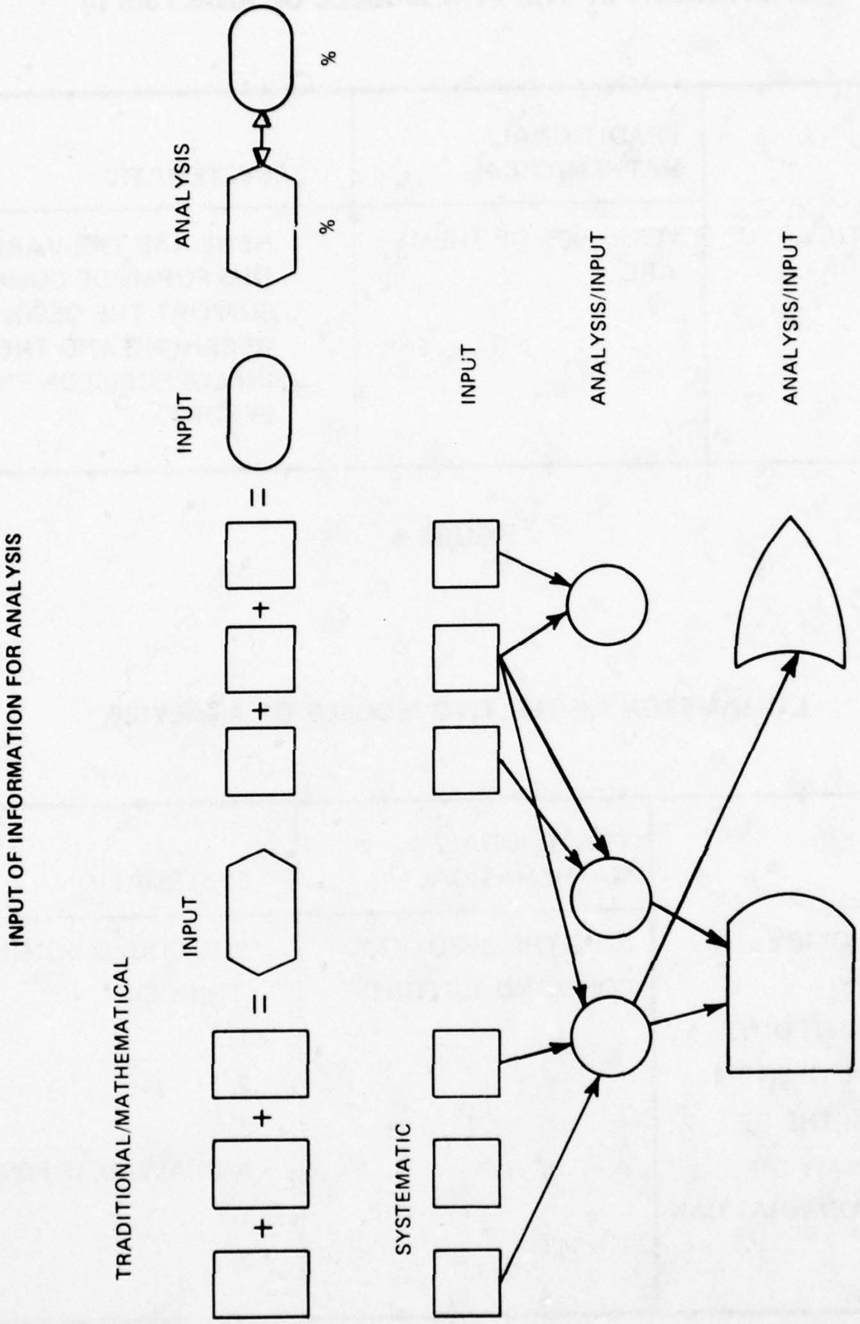


FIGURE 8

The systematic approach upon which our analysis are based, because of its expansionistic nature quickly becomes very complex. As a result data management and analysis is computer dependent and thus incorporated the needed large system computer modeling procedures. How this was done and plans for the future will be outlined by CPT William Nuffer in *Evaluation of the Organizational Effectiveness Effort in the U.S. Army: Implementation.*

QUALITY OF AIR FORCE LIFE:
A REPORT ON THE ATTITUDES AND PERCEPTIONS OF AIR FORCE COMMANDERS

Charles W. McNichols
T. Roger Manley
Michael J. Stahl
Air Force Institute of Technology
Wright-Patterson Air Force Base, Ohio

ABSTRACT

A 1977 Survey af all United States Air Force commanders resulted in 2695 completed questionnaires. The survey was structured around the nine factor quality of Air Force Life (QOAFL) model. Commander's perceptions of their job and their feelings about such issues as quality of Air Force leadership, communications, human relations, pay and benefits, and discipline issues were anonymously solicited in the survey. This report summarizes the results of the survey with emphasis on comparisons across demographic categories.

Introduction

In the Spring of 1975, the Air Force Management Improvement Group (AFMIG) initiated a sequence of attitude and opinion surveys using, as a theoretical framework, a nine-factor model for measuring the Quality of Air Force Life (QOAFL). After the study group was disestablished in the Summer of 1975, the authors continued to assist the Human Resources Development Directorate (AF DPXMMH) in structuring and analyzing QOAFL surveys of various Air Foıce populations. Thus far, a data base of over 50,000 responses has been accumulated from these efforts. Results of analyses of these data have been reported in a variety of proceedings, (Manley, Gregory and McNichols 1975; McNichols, Manley and Gregory 1976; McNichols, Manley and Stahl 1977) and Air Force Institute of Technology (AFIT) theses. In this paper, results of a voluntary, anonymous survey of all Air Force commanders in grades 0-1 through 0-6, distributed in December 1976 will be summarized. Because of space limitations, only a very general summary of survey findings can be presented here. A more comprehensive report is, however, available from Defense Documentation Center (DDC) (Manley, McNichols and Stahl, 1977).

The Sample

Two thousand six hundred ninety-five questionnaires were returned, representing an 80% response rate from all 0-1s thrrough 0-6s with either a commander's AFSC or an A (commander's) prefix to some other AFSC. Response distributions in some major demographic categories are as follows:

Command:	SAC, TAC, MAC, and ATC accounted for almost 70% of the responses.
Grade:	Over half of the sample was serving in the grade of 0-5 or 0-6.
Aero Rating:	47% of the respondents were nonrated, 39% were pilots, and 13% were navigators.

Results

In the following paragraphs, results from major segments of the questionnaire are summarized by topic.

The Commander's Job

The vast majority of respondents (93%) reported that they wanted their present jobs as commanders. A somewhat smaller majority (82%) reported that their present job is challenging, that they have sufficient authority to carry out their responsibilities (77%), and that their present grade is about right for the work that they are now doing (65%).

The 40-hour work week appears to be virtually unknown to the Air Force commander, with almost one-quarter of the respondents indicating that their average exceeds 60 hours. Further, the higher the grade of the respondent, the more hours he she is likely to be working. Commanders of flying organizations and security police commanders report spending the greatest numbers of hours on the job, with about half of both groups indicating that their work weeks exceed 60 hours. Commanders report spending a substantial portion of their time on staff-functional duties and participating in council committee activities, as opposed to "commander's duties." More than one-third of the respondents expressed the opinion that the Air Force requires them to participate in too many activities that are not related to their jobs.

Approximately three-fourths of the respondents expressed the opinion that previous assignments had prepared them for their present commander jobs. However, almost the same proportion indicated that some sort of commanders' training course should be a prerequisite for an assignment as a commander.

Job satisfaction scores, as measured by Hoppock general job satisfaction blank, were higher than those scores obtained from the 1975 Air Force Management Improvement Group (AFMIG) Air Force-wide sample of officers. Lieutenant colonels reported the highest job satisfaction, while lieutenants had the lowest. Similarly, commanders of flying organizations reported the highest job satisfaction, while commanders of services organizations reported the lowest.

Leadership and Communication

Commanders were more positive in their perceptions of both the overall quality of leadership in the Air Force and the leadership ability of their immediate supervisors than was a more representative 1975 sample of the Air Force-wide population of officers.

When compared with the same 1975 data base, commanders report being given greater freedom in doing their jobs, yet they also indicate that they receive recognition less frequently for outstanding work.

While some comparisons with the 1975 AFMIG data reflect positive movement in the area of communication, other comparisons indicate negative trends. Commanders, for example, report that they receive feedback less frequently on job performance and were less likely to receive recognition from supervisors for a job well done.

Military Pay

Commanders perceived that military pay raises are not keeping up with increases in the cost of living, and that their military pay (including benefits) is less than that paid civilians for doing similar work. In comparing commander responses to the questions dealing with pay raises and increases in the cost of living, commander responses were almost identical to those of airmen who participated in the AFIT Military Unionization survey in the Spring 1976. The pattern of commander responses was substantially more negative than those of the officers who participated in the AFIT study, perhaps indicating that the second consecutive "sacrificial" pay raise may have taken a toll among officer personnel.

Career Decision Factors

"My Air Force Job" was selected most frequently by all but lieutenants as having had the most influence on commanders' original career decisions. Lieutenants selected "Training/Education Opportunity" most frequently, with "Security of Air Force Life" and "Retirement System" as second and third respectively. The patriotic institutional theme, "Opportunity to serve my country," was the sixth most frequently cited factor overall. It was placed in that position by all grade groups except for colonels, who had it ranked second.

Substantial differences were noted both in the relative ranking and the frequency of selection, when commander responses to original vs present career decision factors were compared. "My Air Force Job" was selected most frequently by all grade groups when they were asked which factor *today* would influence them the most to make the Air Force a career. While "Retirement System" did not change its second position in the ranking of factors, it did register the single largest increase in percent selected. Also, five percent fewer commanders selected "Opportunity to Serve My Country," dropping that factor from sixth to eighth place overall. It was rated highest (6th) by colonels, and lowest (9th out of 9) by lieutenants.

"Family Separation" was most frequently selected as the factor which would most influence commanders *not* to make the Air Force a career today. Almost 25% of the commanders selected one of two factors dealing with promotion, placing them second and third in the ranking of negative factors. "Fringe Benefits" was selected as the second ranked factor by colonels, but placed sixth by the overall sample. "Pay and Allowances" was placed last by all grades except colonels and lieutenants. Colonels ranked it eighth (out of 10) and lieutenants ranked it ninth. The data in this section seemed to support the Institution-Occupation thesis developed by Professor Moskos.

Discipline, Standards and Enforcement

Commanders considered discipline in the Air Force to be more lax than did AFMIG officer respondents in 1975. A comparison of AFMIG base commander responses with those of base commander today, however, indicates that there has been noticeable improvement.

Commanders are much more positive about the quality of airmen entering the Air Force than were AFMIG officer respondents in 1975. A total of 62a of the commanders stated that the quality of new airmen had improved, while only 12% perceived that it had decreased.

Slightly more than half of the commanders were of the opinion that new airmen arriving from Basic Military Training or Technical Training are motivated to comply with the requirements of Air Force Discipline and Standards. One-third of the commanders held the opposing view.

One segment of the survey presented respondents with a list of subjects for which standards have been established. Opinions of the strictness of standards and enforcement were elicited. Of the 23 standards presented to the commanders, some were perceived to be too strict, some too lax, and others about right. The enforcement of standards, however, is another matter. Using a methodology of measurement developed specifically for this study, only one standard, living in on-base dormitories, was identified as being too strictly enforced. The enforcement of seven other standards could reasonably be described as being "about right," while the enforcement of the remaining 15 standards was characterized as being too lax. Enforcement of standards dealing with overall appearance, wear of the uniform, and military customs and courtesy were considered to be the most lax.

Air Force commanders do not view "standards" as a single, global entity. Rather they tend to discriminate and separate standards into distinct groupings. The same holds true for the enforcement of standards.

The perceived disparity reported by commanders between the appropriateness of standards and their enforcement suggests that commanders are experiencing dissonance. An enforcement/standards dissonance "thermometer" was developed which identified haircuts, mustaches, overall personal appearance, wear of the uniform, military customs and courtesy, and personal weight control as the standards over which the greatest dissonance is experienced.

Overall, commanders perceived discipline in the Air Force to be too lax. Their perception of Air Force discipline, however, was at best weakly associated with the appropriateness of standards. A somewhat stronger association, which might be classified as moderate, was found to exist between commanders' perceptions of discipline and the level of enforcement of *some* standards.

Quality of Air Force Life (QOAFL)

An underlying theoretical model of this research and that of other recent survey research efforts conducted by the staff of the Deputy Chief of Staff Personnel is the nine factor QOAFL model. This model was developed for the Air Force Management Improvement Group effort by the researchers who conducted this study.

Overall, commanders in this study reported higher satisfaction with the quality of their Air Force lives than did the AFMIG respondents in 1975. The factor FREE TIME was the only one for which AFMIG respondents expressed greater satisfaction than did the commanders. Although the rank-orderings of the nine factors by the two groups were found to be highly correlated, the factor LEADERSHIP SUPERVISION warrants mention. AFMIG officer respondents ranked this factor eighth, whereas commanders ranked it third. WORK and HEALTH were ranked first and second respectively by the commanders.

A comparison of the QOAFL responses of the base commanders with those of the AFMIG base commander respondents shows the relative rankings of the nine factors to be very high correlated; strongly indicating that the relative importance of the nine facotrs to base commanders has not changed substantially over the past 18 months. However, today's base commanders expressed greater *dissatisfaction* with eight of the nine QOAFL factors. The only exception was WORK, and that was a tie insofar as dissatisfaction with it was concerned. The largest decreases in *satisfaction* were associated with PERSONAL STANDING, PERSONAL GROWTH, WORK, and ECONOMIC STANDARD. Three increases in satisfaction were noted for HEALTH, LEADERSHIP SUPERVISION, and EQUITY. However, the magnitude of the increases were substantially less than those of the decreases.

Satisfaction with both PERSONAL STANDING and PERSONAL GROWTH among all commanders was found to increase with grade. In other demographic comparisons, both women and black commanders were more satisfied with their PERSONAL STANDING than were others. However, women commanders reported somewhat lower satisfaction with PERSONAL GROWTH than did men.

Job related variables—especially supervisor recognition, being given the freedom needed to do the job well, and challenging work—were found to be strongly associated with both PERSONAL GROWTH and WORK satisfaction.

Commanders' Comments

Commanders were asked to provide comments on any subject which they believed would be of value to Headquarters USAF personnel in their efforts to improve the quality of Air Force Life. Of the 2695 commanders who responded to the survey, 753 (285) provided written comments. This response rate is about double that which is normally received when such a request is made.

While the commanders survey elicited responses in a wide variety of topical areas, the comments section suggests that erosion of benefits tops the list of commanders' concerns, both from the standpoint of the respondents' personal well being and in terms of morale and motivation of the Air Force as a whole.

Programs which deal with the personal appearance of Air Force members received a considerable number of comments. The weight-control program appeared to be a particular sore spot with commanders. More emphasis on rehabilitation and reduced severity of standards were most often suggested. Haircut mustache standards also received considerable attention in the comments section, with most recommendations being for a relaxing of standards, especially for airmen.

The Air Force Officer Effectiveness Report (OER) system (which has undergone major revision since the survey was accomplished) received many written comments in spite of the fact that no specific questions about the system were included in the survey. Almost all commenters were critical of the system for its negative impact on motivation and morale.

Conclusion

The high response rate to this voluntary survey, and the substantial number of written comments obtained, suggest that Air Force commanders were sincerely interested in making their attitudes and opinions known. We are consequently inclined to believe that the survey results represent an accurate reflection of these attitudes and opinions.

REFERENCES

Manley, T.R., R.A. Gregory, and C.W. McNichols, "Quality of Life in the U.S. Air Force, *Proceedings*, 1975 Military Testing Association Conference, Indianapolis, Indiana, September 1975.

Manley, T.R., C.W. McNichols, and M.J. Stahl, Quality of Air Force Life: A Report on the Attitudes and Perceptions of Air Force Commanders, AFIT TR 77-2, Department of Systems Management, School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, Ohio, April 1977 (DDC AD #A042208).

McNichols, C.W., T.R. Manley and R.A. Gregory, "Measuring the Quality of Life of Air Force Personnel," *Proceedings*, 1976 Psychology in the Air Force Symposium, USAF Academy, Colorado, 8-10 April 1976.

McNichols, C.W., T.R. Manley and R.A. Gregory, "Measuring the Quality of Life of Air Force Personnel," *Proceedings*, 1977 Military Testing Association Conference, San Antonio, Texas: October, 1977.

EVALUATION OF THE ORGANIZATIONAL EFFECTIVENESS EFFORT IN THE U. S. ARMY IMPLEMENTATION

William Nuffer
U.S. Army Organizational Effectiveness Training Center

Abstract

Organizational Effectiveness is the Army's program of providing internal Organizational Development consulting services to Army organizations. The program for evaluating organizational effectiveness consists of five phases which are designed to parallel the states of the Organizational Effectiveness Effort — acceptance, training, management, techniques, and cost benefits. The practical aspects of implementing this evaluation effort are discussed; past work and future plans are outlined to include the nature and effects of the system interfaces, questionnaire and interview design, and computer processing details and requirements.

INTRODUCTION

Organizational Effectiveness (OE) is the Army's program of providing internal Organizational Development consulting services to Army organizations. For such a wide-ranging program to be useful, policy makers at all levels need up-to-date and accurate information on the state and progress of the program and advice on the most likely avenues to success. A five-phase evaluation program has been theoretically outlined (Popov, 1978). Accomplishing this evaluation effort in a practical format requires a symbiosis of the behavioral and technological sciences. Data must be collected from as broad an OE involved population as possible with as large a sample size as possible. Because of the interdependence and wide variety of data thus collected, it becomes necessary to use a complex modeling algorithm to provide menaingful results.

DATA COLLECTION INSTRUMENT DESIGN:

The instruments of data collection for each phase consist of one or more of the following:

1. Questionnaires

2. Interview Guides

3. Observation Protocols

4. Records & Reports Reducers

Prior to actual design of the instruments, determination must be made as to what population groups will serve as data collection targets and to what extent they will be sampled.

Data targets for Phases I & II (Figure 1) were initially selected based upon their degree of involvement in the OE system and based upon the potential impact of OE upon the target grouping.

Maximum sampling criteria (Figure 2) was then determined based upon estimated data diversity within each target group (which normally results in a concomitant ease of extraction of the required data elements from the data sources within the group), and upon the logistics and personnel resources available to the researchers.

The goal of this evaluation program being the elucidation of the elements of a complex computer-based model, it was necessary to gather as much data as possible during Phases I & II. The design process for the instruments themselves (Figure 3) reflects this complexity and the top-down systems approach used.

The model, with its elements, areas, and phase goals has been outlined (Popov, 1978). Given the model it is necessary to first determine the specific issues of concern and from this determine next which target group(s) are most appropriate to provide the information required and thence to construct the specific means (Figure 3) of data extraction. It is interesting to note that the data gathering process becomes less general and more specific as we progress thru the phases. Data from previous phases allow us to increasingly narrow our focus.

A difficulty in obtaining the data, once the tools are available, is the climate of fear which leads many persons in positions of responsibility to believe that any information gathered will be used against them. Particularly in Phase I, it was necessary to assure commanders and senior staff of not only individual anonymity but of organizational anonymity as well.

DATA HANDLING:

This method of evaluation research does not embrace "data analysis"; rather, the data handling process is a method of system analysis and data collection, followed by "data synthesis". Data synthesis takes several forms, the most crucial of which is the OE system model. Of interest also are two special applications which perform key roles in this evaluation.

Traditionally the computer has been merely a tool of statistical manipulation for the behavioral science researcher. Not until the introduction of System Dynamics (Forrester, 1961) was the computer put to any creative use in the behavioral science milieu. Later application in the urban (Forrester, 1969) and world (Meadows, 1972) social environments proved the use of the computer as a tool for social research. Utilizing System Dynamics our GOET (Growth of OE Over Time) model explores the dynamic interdependency of OE demand, supply, and credibility.

The credo of OE implicitly purports to increase "effectiveness" (more specifically "mission" or "combat" "effectiveness"), yet there is no common agreement as to what this elusive "effectiveness" is, let alone how to measure it. Another part of our effort (embodied in Phase III) involves measuring effectiveness not only through the traditional (and often misleading) use of command statistics (IG report, readiness reports, test scores, etc.) but through the perceptual data from commanders and subordinates alike. Figure 4 outlines this effectiveness measurement system.

PHASE II DATA TARGETS

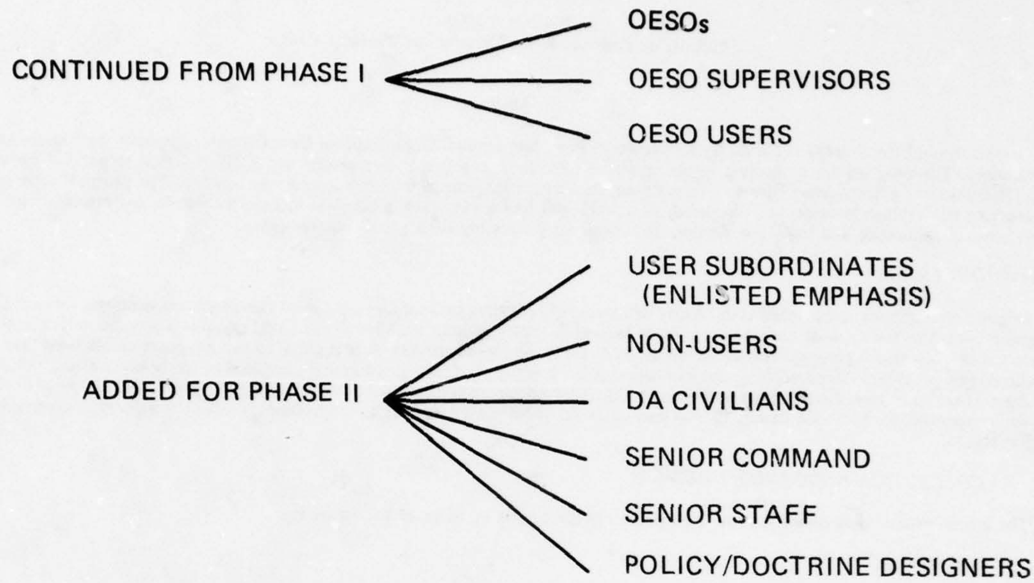


FIGURE 1

PHASE II SAMPLING CRITERIA

OESOs	100%
SUPERVISORS	100%
COMMAND (06-08)	REPRESENTATIVE
SENIOR STAFF (05-07)	REPRESENTATIVE
USERS (02-06 GS9-15)	100% AS SUPPLIED BY OESOs
USER SUBORDIANTES (E4-02)	10 PER USER
NON-USER (02-06)	REPRESENTATIVE AT SELECTED LOCATIONS
CIVILIAN (GS9-15)	50% AT SELECTED LOCATIONS

FIGURE 2

AD-A068 221 AIR FORCE ACADEMY COLO F/G 5/10

PROCEEDINGS OF THE ANNUAL SYMPOSIUM ON PSYCHOLOGY IN THE DEPART--ETC(U)

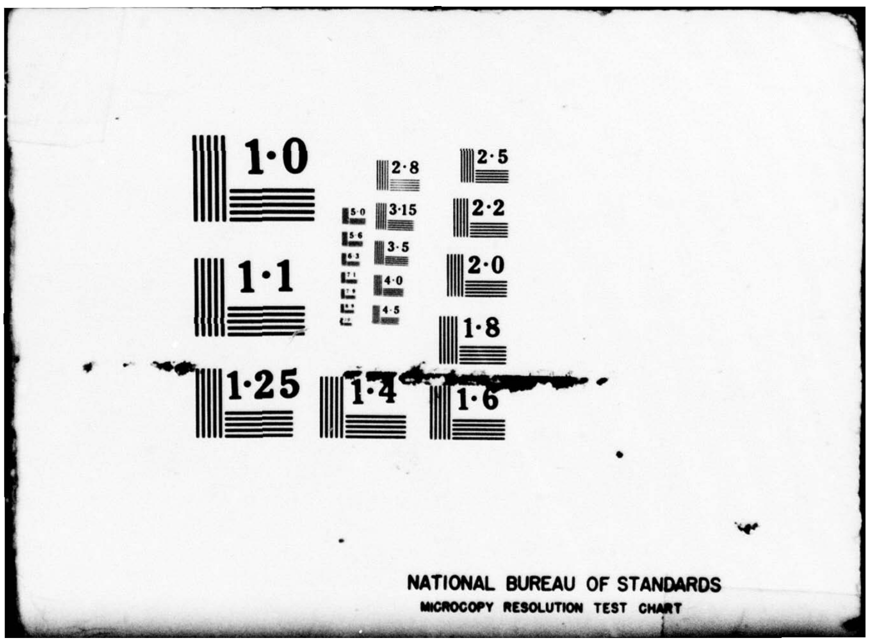
APR 78 J M BERMUDEZ, W H CLOVER, L J DAHLE

UNCLASSIFIED USAFA-TR-76-12 NL

2 OF 4

ADA

068221


1·0
1·1
1·25
1·4
1·6
1·8
2·0
2·2
2·5
2·8
3·15
3·5
4·0
4·5
5·0
5·6
NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

INSTRUMENT DESIGN PROCESS MATRIX

GIVEN: 1. Phase Goals
2. Results of previous phases
3. Target Selection

Questionnaire Matrix

Targets

Model Element	Element Area	Area Issue	. . . .	OESO	. . . .
.	.	.			
.	.	.			
	.	.			
		.			
		.			
ORG Climate	Personal Support System	1. What, if any, characteristics of the OESO predispose satisfaction with the OESO role?		How satisfied are you with your role as an OESO? (Linked to characteristics questions)	
		.			
	.	.			
.	.	.			
.	.	.			
.	.	.			
.	.	.			

FIGURE 3

OVERVIEW OF PHASE III MISSION EFFECTIVENESS EVALUATION PROCEDURE

Types of Units: Combat Brigades
Support Cmds
Cbt Support Bns
Training Brigades
Civilian Heavy Support Activities

Flow of Collection and Analysis:

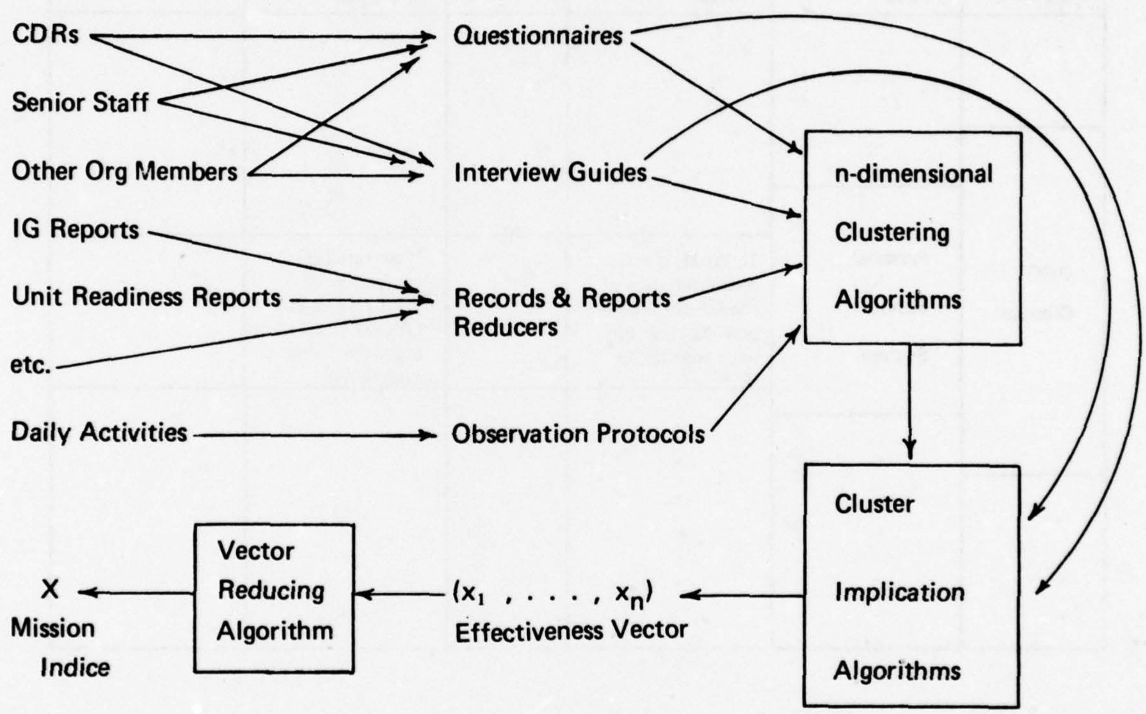


FIGURE 4

Last and most important the computer model of the OE process and its interfaces is proposed and is currently in the middle design phase. To rationally model the overwhelming complexity of the OE system requires new concepts, borrowing the techniques of pattern recognition from the system sciences and list-processing from the field of Artificial Intelligence.

It becomes possible to overlay a recursively pointer-linked template upon the elements and areas of the model. Then by executing a finite limit recursive pattern search, an *n*-tensor of order 2 can be constructed where n is the number of pattern elements. This tensor then becomes amenable to standard statistical procedures, principally involving cluster analysis.

DISCUSSION:

New and controversial techniques of evaluation are being applied in hopes of providing a realistic feedback system to policy makers. The ulitmate fate of the evaluated system (OE) is dependent upon the type, accuracy, and timeliness of the data provided. A following paper (Best, 1978) exemplifies some of the information already provided through these techniques.

BIBLIOGRAPHY

1. Best, Paul, "Evaluation of the Organizational Effectiveness Effort in the Army: First Results and Their Significance in a Management Focussed Evaluation Effort," in *Sixth Psychology in the DOD Symposium Proceedings*, 1978.

2. Forrester, Jay, *Industrial Dynamics*, MIT Press, 1969.

3. Forrester, Jay, *Urban Dynamics*, MIT Press, 1969.

4. Meadows, Donella, et al, *The Limits to Growth*, Potomoc Associates, 1972.

5. Popov, Dan, "Evaluation of the Organizational Effectiveness Effort in the U.S. Army: Evaluation Design," in *Sixth Psychology in the DOD Symposium Proceedings*, 1978.

EVALUATION OF THE ORGANIZATIONAL EFFECTIVENESS EFFORT IN THE U. S. ARMY FIRST RESULTS AND THEIR SIGNIFICANCE IN A MANAGEMENT FOCUSSED EVALUATION EFFORT

Paul Best
U.S. Army Organizational Effectiveness Training

Abstract

The results of the first two phases of the five-phased Army-wide Organizational Effectiveness (OE) Evaluation are based on 1300 comprehensive surveys and 440 detailed interviews covering the following aspects of OE: the local organizational climate for OE, the important characteristics, knowledge and skills of the Army-trained OE consultants, and the techniques used by the consultants. The data collection and computer modeling procedures used allow for the ability to put together information bearing on almost limitless questions issues raised by policy-makers and practitioners. Selected results demonstrating the overall state of Army OE will be presented and discussed in terms of their significance from a management perspective.

INTRODUCTION:

The unique design and implementation of the Army's Organizational Effectiveness (OD) evaluation effort allows for the policy-makers to use recently collected data for informed decision-making. This capability provides an immence potential bonus to the developers of a new, large-scale program such as Army Organizational Effectiveness (OE).

The purpose of this paper is to present selected results to illustrate the significance of evaluation data for management and decision-making.

RESULTS – SUPPORT FOR OE:

Phase I of the Evaluation Program focussed on the acceptance of OE at the installations where it existed during the first year of the Army OE program (Jan-Dec 1976). One of the critical elements of the organizational climate portion of the evaluation model is local command and staff support for the consultant and his/her efforts. Command Staff support obtained from the surveys and interviews produced five distinct support types which are described in Figure 1. The first three support types shown in Figure 1 are associated with accepted programs, but "Carte Blanche" and "Laissez-Faire" are associated with the optimally accepted programs. These two support types also represent the support most desired by the consultants.

TYPES OF SUPPORT

CARTE BLANCHE	Support in terms of behavior attitudes and leadership style of command. OESO has direct access to commander, good resource support.
LAISSEZ FAIRE	OESO left alone and not helped nor hindered. OE users have implicit permission to use OE.
GUARDED SKEPTICISM	Command staff are skeptical. Wait and see attitude.
POLITICALLY ADVANTAGEOUS	Support is primarily lip service OE is viewed as a political necessity.
NO SUPPORT	No contact with commander, no staff support. Poor resource support.

FIGURE 1. DESCRIPTIONS OF COMMAND STAFF SUPPORT

The Phase I data indicated that more emphasis could be placed on appropriate command support from Department of Army. Due to the lack of policy and guidance in this area each local installation developed its own unique style of support based on the individual personalities of people in key positions. The descriptions of command staff support produced by the consultants, their supervisors and the command staff differed significantly indicating that each of these groups viewed the support issue differently further delineating the complexity of this issue.

The extent of each support type at the end of December 1976 (Phase I) and again eight months later (Phase III) is shown in Figure 2. These data indicate that "Carte Blanche" and "Laissez-Faire" support both increased over the eight month period while the remaining three types of support decreased. During this eight month period the Army OE program was further developed—more consultants were supplied to the field, various educational programs were conducted and DA guidance was published.

The importance of the issue of command support was underscored by the finding that the type of support had a major impact in three key areas: (1) Accceptance of OE by the local population; (2) the consultants' personal satisfaction with their role and decision to join the program; and (3) the resource support provided to the consultant to accomplish his her job. These are all very important areas for a developing program. The descriptive model provided information on the current state of the program, the impact of this state on the future of the program and produced prescriptions to alter it to a more desirable state.

RESULTS – TRAINING, SELECTION AND ASSIGNMENT OF OE STAFF OFFICERS:

The primary focus of Phase II of the evaluation effort was to determine the appropriate training, selection and assignment criteria for the OE Staff Officers. The evaluation model was designed so that information on acceptance was obtained first followed by information on how to fine-tune training, selection and assignment. This provides the capability to match the OE Staff Officers (based on background and OE skill) with the assignment location so as to maximize the likelihood of accomplishing the goals of the OE program.

The surveys and interviews obtained information from all respondents as to the types of training the OE Staff Officers should have, additions and deletions to the training curriculum, characteristics of OE Staff Officer and their jobs that lead to satisfaction or dissatisfaction, etc. Additionally, information was obtained as to the personal characteristics, military experience, military and civilian education, etc., they should have before an OE assignment.

Overall we found that the ideal OE Staff Officer differed between the military and civilian respondents. Whereas, the military respondents desired a former commander with behavioral science education at the 0-4 rank, on active duty, the civilians perferred a person with business administration background, also a former commander, but retired preferably at the 0-5 level. The characteristics are presented in decreasing order of importance in Figure 3. These results provide an idea as to how to maximize the potential of OE at an installation: send the person with the right characteristics so as to predispose support, credibility and confidence from the population at hand.

When the military data is partitioned according to the level of the unit, i.e., battalion and smaller, major unit and major command, there are unique requirements for OE Staff Officers at each level. For example, the OE Staff Officer who will work primarily at the battalion level and below should be a senior 0-3, with command experience and a civilian educational background in behavioral sciences. When the focus of operations is at the major unit level the OE Staff Officer should have both command and staff experience, be an 0-4 with educational experience in the business management area. They should also be a member of the same branch to which the unit belongs, e.g., infantry officer with an infantry unit. These criteria for selection and assignment are listed in Figure 4, which also depicts the criteria at major command level and for civilian-heavy organizations. Taken all togehter these results provide criteria by which to select and assign OE Staff Officers for maximal impact within their assigned unit.

Finally, the data provided information on the appropriate training areas for each of the levels of focus discussed above. At the battalion level and below the OE officer must be knowledgeable in assessing the general climate of the organization, working on interpersonal skills and assisting in transition of command. At the higher levels they must be prepared to work on large organizational systems, inter-staff relations and the particular technology of the unit to which they are assigned, e.g., budgeting, personnel management, etc. These training emphasis results are presented in Figure 5. The results would allow for the development of a course of instruction which includes a core curriculum followed by assignmnet-specific training. Such a training program would be the most efficient method for obtaining maximal impact toward the OE goals in the assigned organizations.

SUMMARY AND CONCLUSIONS:

Examples of the use of evaluation data to provide feedback to aid in the development of a large-scale Army program have been presented. These data are part of a large-scale, logic-based computer model which will allow assessment of many varied aspects of the program as the need arises as well as to conduct the goals of the evaluation program itself. There are many reasons for implementing new programs or modifying on-going programs and often assessment of the needs of the field and key personnel does not provide all of the pertinent information. However, the evaluation program described in this report and by Popov and Nuffer (contained in this proceedings) is a viable way to present maximal, timely information during the development of a program as well as to collect necessary data for eventual program evaluation.

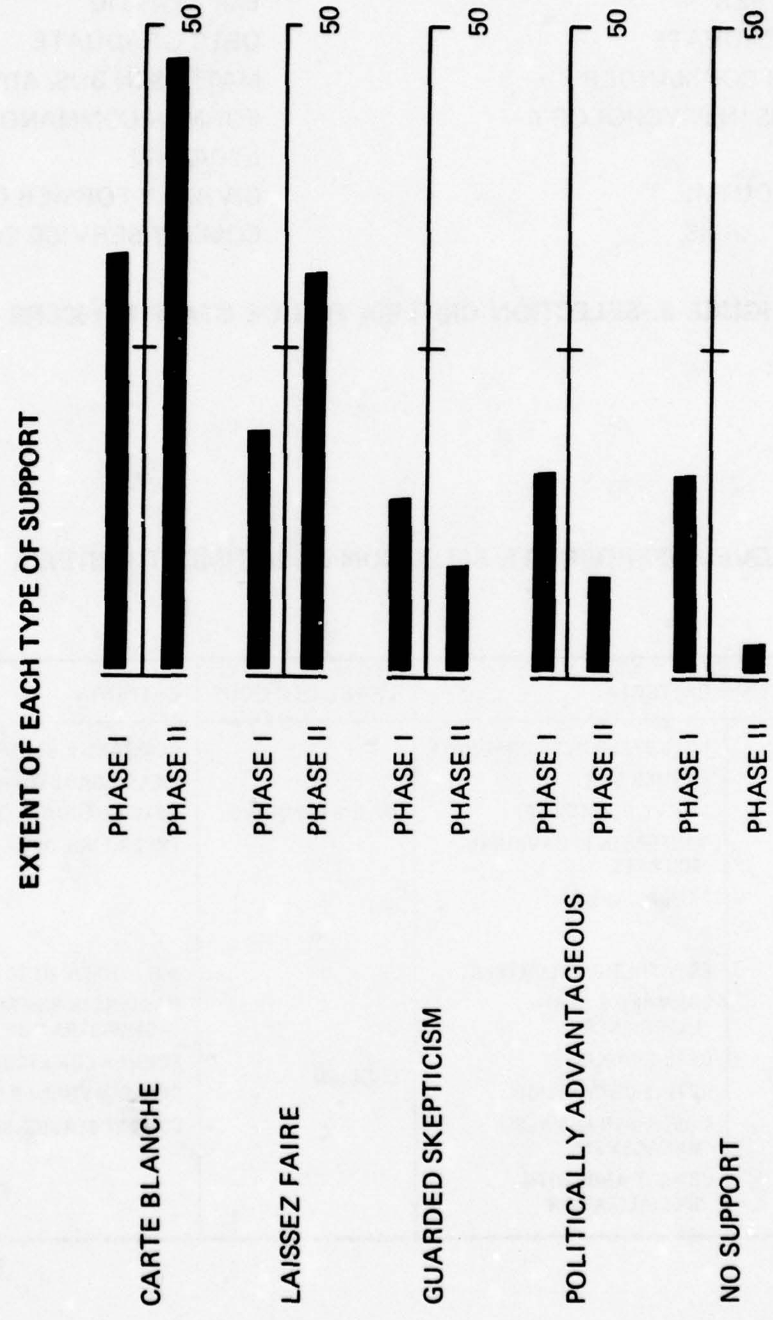


FIGURE 2. DISTRIBUTION OF SUPPORT TYPES — PHASE I AND PHASE II

SELECTION CRITERIA FOR OPTIMAL CREDIBILITY

MILITARY POPULATION	CIVILIAN POPULATION
VOLUNTEER	EMPATHETIC
OETC GRADUATE	OETC GRADUATE
FORMER COMMANDER	MATERS IN BUS. ADMIN.
MASTERS IN PSYCHOLOGY	FORMER COMMANDER
MAJOR	LTC/GS-12
ACTIVE DUTY	CIVILIAN-FORMER OFFICER
COMBAT ARMS	COMBAT SERVICE SUPPORT

FIGURE 3. SELECTION CRITERIA FOR OE STAFF OFFICERS

LEVEL APPROPRIATE SELECTION ASSIGNMENT CRITERIA

LEVEL OF FOCUS	CRITERIA	LEVEL OF FOCUS	CRITERIA
BATALLION AND BELOW	VOLUNTEER OETC GRADUATE FORMER CDR ACTIVE DUTY SR CPT MASTERS IN BEHAVIORAL SCIENCES COMBAT ARMS	MAJOR COMMAND	COMMAND & STAFF EXPERIENCE FIELD GRADE WITH C&GS OETC GRADUATE COMBAT ARMS EXPERIENCE AS OESO MATURITY
MAJOR LINE UNIT	EMPATHETIC VOLUNTEER COMMAND & STAFF EXPERIENCE OETC GRADUATE ACTIVE DUTY MAJOR MASTERS IN BUSINESS/ MANAGEMENT COMBAT ARMS WITH SPECIALIZATION	CIVILIAN	EMPATHETIC OETC GRAD MASTERS IN BUSINESS ADMINISTRATION FORMER CDR LTC/GS12 CIVILIAN FORMER OFFICER COMBAT SERVICE SUPPORT ARMS

FIGURE 4. SELECTION ASSIGNMENT CRITERIA AS A FUNCTION OF LEVEL OF FOCUS

LEVEL APPROPRIATE TRAINING EMPHASIS

LEVEL OF FOCUS	CRITERIA	LEVEL OF FOUCS	EMPHASIS
	ASSESSMENT OF ORGANIZATIONAL CLIMATE		ASSESSMENT OF LARGE SYSTEMS
			MOTIVATION
	INTERPERSONAL SKILLS		SYSTEMS DESIGN
		MAJOR COMMAND	
BATTALION AND BELOW	COMMAND TRANSITION		PLANNING
	MISSION ANALYSIS		OPERATIONS RESEARCH TECHNOLOGY
	TRAINNIG OBJECTIVES		DECISION MAKING AND MANAGEMENT
	LEADERSHIP AND MANAGE-MENT DEVELOPMENT		
	ASSESSMENT OF INTER-SELF RELATIONS		CIVILIAN-MILITARY INTERFACE
	TEAM BUILDING		JOB SATISFACTION/ ENRICHMENT
	FORMAL ORGANIZATION STRUCTURE		CIVIL SERVICE SYSTEM
MAJOR LINE UNIT		CIVILIAN	
	MANAGEMENT BY OBJECTIVES		MANAGEMENT BY OBJECTIVES
	ORGANIZATION APPROPRIATE TECHNOLOGY (BUDGETING, PERSONNEL MANAGEMENT)		ORGANIZATION APPROPRIATE TECHNOLOGY
	MANAGEMENT TECHNOLOGY		BUSINESS ADMINISTRATION

FIGURE 5. TRAINING EMPHASIS AS A FUNCTION OF LEVEL FOCUS

RELATION BETWEEN LEADER BEHAVIOR, ORGANIZATIONAL CLIMATE, AND UNIT EFFECTIVENESS

Hal W. Hendrick
University of Southern California

During the past three years I served as the head of a Defense Department organizational assessment team comprised of faculty members from the Defense Race Relations Institute. In most cases, we were called in to do assessments of large military units which had been experiencing a number of racial incidents. In assessing these organizations, we found that most of the actual problems were not purely racial and that other organizational climate dimensions also were affected. The primary factor which consistently showed up as being strongly related to the organizational climate dimensions was the functional behavior of the company level commander. While the behavior of superior commanders clearly had some influence on the company commanders, and although subordinate leaders' behavior was influenced by that of the company commander, no strong relation of their behavior to the climate and to the effectiveness of individual companies was found. It appeared from our assessments that the strong relationship was found for company commanders because the company is the lowest level or relatively autonomous functioning unit within the spectrum of military units assessed. In addition, the company commander has strong positional power and fairly frequent direct contact with all subordinates.

The above described findings led us to do additional assessments focused at the company unit level. From these assessments a model of the relationship between subordinate perceptions of leader behavior, subordinate feelings and motivation, organizational climate, and unit effectiveness emerged. The purpose of this paper is to describe this empirically derived model.

METHOD

Subjects

The subjects used in developing this model were 400 male enlisted personnel from 60 Army companies located on two large posts in the south-central United States. Included in the sample were infantry, artillery, armored, supply, and military police companies. Six to eight personnel from each company were interviewed. One half of the sample were minority group members.

Procedure

Each subject included in the study was interviewed by a member of the assessment team. Interviews were conducted at the subject's place of work, in the barracks, or in a service club. The semi-structured interview was designed to elicit the subject's perceptions of the (a) company commander, (b) various organizational climate dimensions, (c) unit effectiveness, and (d) personal feelings and attitudes about the subject's role in the company.

Based upon a content analysis of the subjects' responses to questions about organizational climate factors and unit effectiveness, the subjects' companies were categorized into four criterion groups as shown in Table 1. For all of the 54 units shown, there was a uniformly high consistency among the criterion evaluations by subjects from the same company. In the six additional companies results were not consistent across subjects for at least one of the criterion dimensions. These six companies were dropped from the study. As shown in Table 1, none of the companies were evaluated as both having a healthy organizational climate and being below average in unit effectiveness.

TABLE 1
Classification of Companies by Perceived Organizational Climate and Unit Effectiveness

Organizational Climate	Unit Effectiveness	
	Above Average	Below Average
Healthy	19	0
Poor	7	28

RESULTS

Based upon a content analysis of the interviews, two functional categories of leader behavior were identified. These were the familiar dimensions of *initiation structure* and *consideration* frequently found in other studies of leadership (Stogdill, 1974). However, results of the present study identified possible refinements to our present knowledge of the consideration dimension. Specifically, three facits of consideration emerged: Demonstration of concern for subordinates welfare; not showing favoritism; and the nature of leader-subordinate transactions--in T.A. terms, were the transcations adult-adult or parent-child in nature. Another aspect of the transaction facit emerged only in interviews with minority members, but it emerged consistently; this component was the extent to which the leader demonstrates respect for subordinate differences in racial cultural values, life styles and mores. High consideration leaders were described as interacting on an adult-adult basis with subordinates, respecting racial cultural differences, showing concern for subordinate problems and welfare, and not showing favortism. When asked for their global perception of the leader's consideration behavior, it appeared that the transaction facit carried the most influence, but with two notable exceptions: When the leader was seen as practicing blatant favoritism, and when he was seen as going out of his way in demonstrating true concern for subordinate problems and frustrations. In short, favoritism functioned like a hygenic factor, and demonstrating concern like a motivator in influencing global perceptions of leader consideration. The frequency of occurance of each leader behavior pattern within the four criterion groups is shown in Table 2.

TABLE 2
Relation of Perceptions of Leader Behavior to Organizational Climate and Unit Effectiveness

Leader Behavior	Criterion Group		
	Healthy Clim High Effect	Poor Clim High Effect	Poor Clim Low Effect
High Struct. High Consid.	16		
High Struct. Low Consid.	3	7	23
Low Struct. High Consid.			2
Low Struct. Low Consid.			3

As may be noted, leaders perceived as high on both structure and consideration behavior also were perceived as having effective companies with a healthy organizational climate. Only one out of three leaders with high structure and low consideration were perceived to have effective organizations by their subordinates, and less than one in ten as having a healthy organizational climate within their companies. Further analysis of the interviews yielded the following model of the relationship between leader behavior, subordinate feelings and motivation, organizational climate dimensions, and unit effectiveness.

Figure 1 depicts the model for an organization headed by a leader who exhibits high structure, low consideration behavior. Subjects having this type of company commander typically stated that all the commander cared about was accomplishing assigned tasks and staying out of trouble with the "old man" so as to get a good effectiveness report and get promoted. The leader was seen as being concerned about his own future at the expense of subordinates. The feeling was frequently expressed by subjects that the commander did not care about him as a person, did not respect him as an adult, and did not value his contribution to the unit. When asked how this made him feel, subjects consistently stated they felt frustrated, discouraged; and angry. Most further reported that, as a result, they simply tried to stay out of trouble, get away whenever they could, and do just enough to "get by."

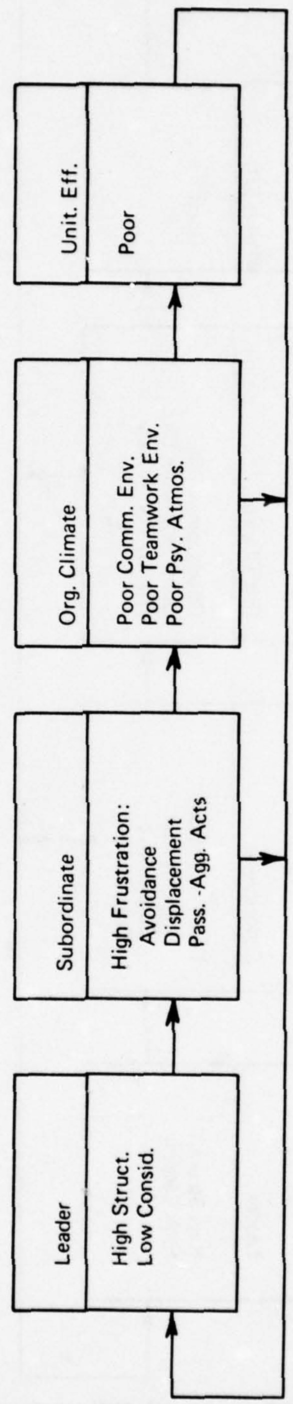


FIGURE 1. RELATIONSHIP FOR LEADER EXHIBITING HIGH INITIATION STRUCTURE AND LOW CONSIDERATION BEHAVIOR

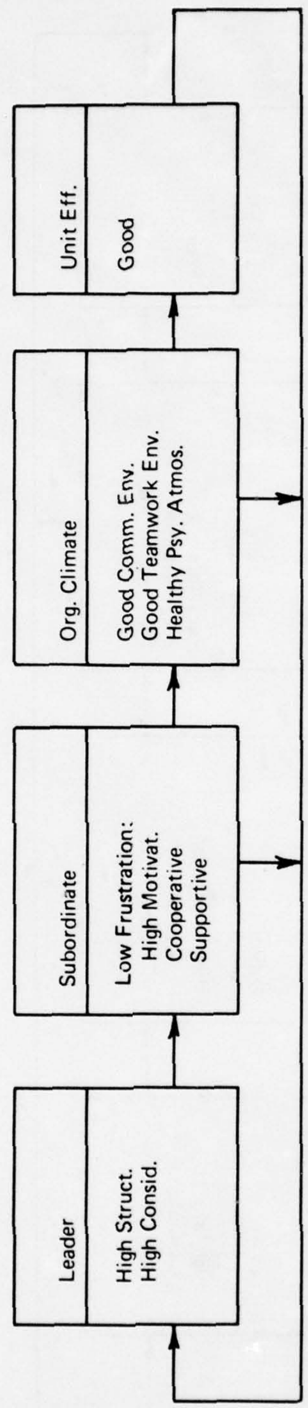


FIGURE 2. RALATIONSHIP FOR LEADER EXHIBITING BOTH HIGH INITIATION STRUCTURE AND HIGH CONSIDERATION BEHAVIOR

This high level of frustration on the part of unit members appeared to be acted out in the form of *displacement* of aggression on to peers and subordinates, and in *passive-aggressive acts* as a means of "getting back" at the "system" in general and the company commander, specifically. In the interviews the displacement was reported in the form of acts of aggression toward team members and others in the organization--particularly those who differed along racial or ethnic lines. It appeared to us that under these conditions of high frustration, deep seated prejudices which normally would not be acted out served as a readily available avenue for displacement. In addition to the disruptions of communications caused by displacement, subjects reported what appeared to be disruptions from passive-aggressive acts: Messages would come to them and they would not listen carefully or make the effort to record them; at times, they would "forget" to relay the messages; often subjects reported "taking their time" in delivering messages and in doing other tasks. In several cases, subjects admitted to having deliberately distorted communications and to "messing up" assigned tasks. The cumulative effect of these acts was reflected in measures of various organizational climate dimensions as shown in Figure 1. Subjects typically saw these personal frustrations and the poor organizational climate as resulting from poor leaderhsip which, in turn, accounted for the perceived poor unit effectiveness.

Figure 2 dipicts the model for the case where the company commander was perceived as being high in both initiation structure and consideration behaviors. This appeared to be the condition in which subordinate frustration was minimal and in which leader induced motivation was maximal. The high initiation structure behavior ensured that subordinates knew organizational goals, their jobs, the expected standards of performance, and organizational policies and procedures. As a result, frustration which can eminate from not knowing what to do, how to do it , or what is expected by way of performance was minimized. When subordinates perceived their commander as respecting them as adults, respecting their race or culture, valuing their contributions to the unit, and genuinely concerned about their welfare, not only was frustration minimized but motivation to "really put out for the old man" was often reflected in the interviews. Subordinates would state that they and the others in the unit would make that extra effort to ensure that communications were received, recorded, and relayed promptly and accurately. Instead of acting out frustration through displacement, subjects reported good relations among team members and within the unit in general. Of particular note was the often stated positive relations among unit members of differing racial and ethnic backgrounds. In every one of the sixteen companies conforming to this case of the model, unit effectiveness was consistently reported as being high.

In examining these first two cases of the model, another factor became apparent to the assessment team: Whatever the *outcome* of the leader's behavior, it served to reinforce continuation of the same behavior on the leader's part. This was true even when the behavior was disfunctional. For example, when high initiation structure and low consideration behavior led to less than optimum unit performance, communications breakdowns, discipline problems, and racial incidents, the commander's response to this increased pressure and personal workload was to further intensify his initiation structure behavior and to exhibit even less consideration behavior. In this manner he tended to perpetuate, rather than reduce, organizational problems. This circular reinforcement of both functional and disfunctional organizational patterns is represented by the feedback arrows in Figures 1 and 2. Although based only on a few cases, the same reinforcement pattern also appears to hold for our other two cases of the model, where the leader is either low in initiation structure and high in consideration, or where he exhibits little of either kind of leadership.

Only two company commanders in the study were rated high in consideration and low in initiation structure behavior. Both of their companies were charcterized by at least some subordinate frustration and a somewhat poor organizational climate. Both units also were seen as below average in effectiveness. In both cases the company commanders, and therefore the companies, were seen as continually in trouble with the superior commander. In addition to the frustration resulting from this situation, failure of the company commander to provide more direction also seemed to have become a major source of frustration to company personnel. The ensuing pattern of displacement and passive-aggressive behavior, poor organizational climate, and reduced unit effectiveness seemed to have resulted, much as in the high initiation structure, low consideration case.

From review of the first three cases of the model, it is easy to understand why the three companies in the fourth case also were perceived as having a poor organizational climate and below average unit effectiveness. Failure to provide structure or to show consideration allowed the absence of both behavioral dimensions of leadership to serve as a source of subordinate frustration.

DISCUSSION

The model empirically developed in this study potentially offers additional insight into the dynamics of organizational functioning, particularly as it is affected by leader behavior. In addition, the study has shed additional light on the dimension of consideration leadership behavior. That these additional components of consideration have more to do with *how* a leader behaves than with engaging in some additional kind of behavioral task is particularly noteworthy, as this finding has direct implications for leadership training and development.

At the present time, the model must be regarded as limited in the extent to which it can be generalized to other kinds of organizations, and to other levels of military organizations. Empirical studies of other organizations and of various levels of organization are needed to assess the validity of the model across different organizational structures, environments and sizes. In the present study the leaders under evaluation had very high positional power. Studies of organizational units where the leader has only limited positional power also are needed to further test the model.

REFERENCES

Stogdill, R.M. *Handbook of leadership*. New York: Macmillan, 1974, 128-141.

THE LEADERSHIP INFLUENCE MATRIX (LIM)

Loren I. Moore
U.S. Navy

Leadership models and research to date have not provided a comprehensive concept which includes the many variables upon which effective, efficient leadership in a given situation may be postulated. The interrelationships of the elements of leadership, the leaders, the followers, and situations are arrayed in the Leadership Influence Matrix (LIM) providing: an overall concept, a technique for organizational diagnosis, a means of program classification, and one basis for appropriate leader behavior.

The need for leaders has been recorded throughout history. The *Bible* even reports history in relation to the leader. Increasing bureaucracy has spawned many words with varying differentiation between meanings and functions, i.e. leadership, leader, management, manager, boss, director, chairperson, administrator, ruler, superior, supervisor, teacher, coach, commander, and the like.

Leadership in the broadest context is the background for this Influence Matrix exploration. Leaders and followers are continually faced with major issues of leadership: responsibility, authority, delegation, goal setting, control, time management, decision making, problem solving, optimum task achievement, communications, performance evaluation, team building, conflict management, etc. There is a host of questions about theory, process, and practice of leadership. How can a leader get the job done most effectively? How does one deal with the myriad of variables? When should one listen and when should one give orders? When and how should one use power? How and when does one reward and punish? How can an organization or an individual put leadership theory and research data into practice?

In the past, leadership models and research did not provide concepts upon which situationally effective leader behavior could be postulated (Hersey and Blanchard 1976; Fiedler 1976). The question remains: how does the leader decide how to act or behave? Perhaps the answer is simpler than one might suppose. It is here posited that appropriate, effective, efficient leadership is situational leadership that provides systmeatic, pragmatic answers to the Who, What, Why, Where, When, and How questions of leaders and followers.

Leadership is influence, influence is power, and "power is essential to all living things" (May, 1971). The emphasis upon hierarchical organizations and models, particularly the family, the church, the military, and business, tends to predispose one to think of power as being possessed and exercised only by those in the superior position. Leadership conceptualized as an interpersonal relationship, shows that all the elements of group effort (leader, followers, and situation) possess power or influence. It may be helpful to think of these influence relationships in a Leadership Influence Matrix (LIM) as shown:

Influence	Leader	Followers	Situation
Leader	Leader Influence on Leader	Leader Influence on Followers	Leader Influence on Situation
Followers	Followers Influence on Leader	Followers Influence on Followers	Followers Influence on Situation
Situation	Situation Influence on Leader	Situation Influence on Followers	Situation Influence on Situation

The elements of the Matrix may be examined in some detail.

Leader Influence. Examples of the leader influence upon leader of the 1930's and 1940's, the personal development aspects of managerial and leadership development programs, and the use of T-groups and individual centered methodologies of the past. The Transactional Analysis (Harris 1967) and Education of Self (Weinstein 1970) approaches to L/L influence offer much promise. Further, emergent areas of third force psychology, trans-personal counseling, and psycho-synthesis have potential in this area of leader self knowledge. The use of Gestalt techniques to obtain "Authentic Management", individual time management, individual approaches to individual goal setting, to individual decision making, and to individual life planning, and the general theory of "improved organization through improved individuals" is indicative of the thrust with this intersection of the Matrix.

Currently one of the most important aspects of L/L influence is the leader's personal racism, and sexism as it is manifested in the leader behaviors.

Leadership development programs that are based exclusively upon L L influences will usually be effective in the short run, (the increase in personal skills and feelings of growth for many participants will normally provide positive data). In the long run an exclusive L/L approach is fated to be less than optimum. On the other hand any leadership development program that ignores L/L will also be less than optimum.

Leadership influence on followers, L/F, is the influence relationship that has received the majority of focus to date. In terms of the Situational Leadership Influence Model (SLIM) (Moore, 1977), this is leader behavior (LB) that represents the behavior of the leader in relationship to the influence development (ID) level of the followers. This leader behavior can be thought of in terms of skills or processes. Skills such as decision-making, problem solving, time management, goal setting, conflict resolution, intergroup negotiation, personnel development, training, that favorite, communications, and the many, many other skills.

Simply stated these are the skills and processes that are manifested as leader *structuring* and *consideration* behaviors. Together these behaviors encompass all of the leaders (and followers) leadership activities. Any program that positively addresses the leaders skills and processes will normally enhance leadership and usually enhance task accomplishment. A leader skills approach to formal leadership training and education has been traditional. As it is a part of the Leadership Influence Matrix, the approach is generally successful. To the extent that either the formal or informal organization does not include the other appropriate elements of leadership, this L/F approach will not optimize leaders and followers potential.

Leader influence upon the situation, L/S, is an area that has not been systematically studied. The leader or a group of leaders may be thought of as the action initiators in terms of Forrester's "counter-intuitive behavior of complex organizations." The thesis seems to hold that leader behavior in large systems may be counter-intuitive; i.e., what happens is not what the leader thinks should happen. (Forrester, 1971). Organization Development (OD) to date still strongly represents its background in psychology and the personal L/L approach to organizations. This personal L/L approach is largely responsible for the wide disparity of OD results in different organizations, or even in the same organization at different times, with the same or different peoples. For true systems OD, the whole leadership Influence Matrix (LIM) must be considered. Most OD efforts, it is held, belong in the leader influence upon leader, L/L or leader influence upon followers L/F, sections of the Matrix.

All elements of the LIM that contain the situations(s) influence have had little research accomplished. Until the advent of the computer, there were simply too many variables in "situations". Even today the situation(s) influences are not chief subjects of investigation.

Follower Influence. The influence of the followers on the followers (F/F) concepts may be seen in role theory and in associated areas. Follower influence on followers, F F, is vital to situational leadership theory in that as the followers have higher levels of influence development, ID the group must exercise the same skills in decision-making, communication, goal setting, et cetera, that the leader formerly did. That these follower skills are required is seen by the emphasis upon team building, communication skills, trust building, group decision making, group goal setting, group creativity, and the like in many leadership, management, and OD programs.

Regarding the followers influence upon the leader F L, one aspect which has been labeled but is relatively uninvestigated is F/L on a one-to-one basis. The other aspect, almost as equally univestigated, considers the followers as a group and can be best states: "What is and how do you determine follower influence?" The major component of follower influence upon the leader can be expressed in influence development, ID, levels. A Situational Leadership Influence Model (SLIM) has been developed (Moore, 1977) that provides insights and techniques in the determination of follower influence.

The influence of the followers (as a group or as individuals) upon the leadership situation, F/S, has not, as is to be expected, been explored in any depth. The mass movements of the past (civil rights, labor, female liberation) might be studied in that light.

Situation Influence. The influence of the situation upon the leader, the followers, and on the situation has been focused in various areas: The OD approach previously mentioned, Fiedler's approach, and the "analysis" approach. Thompson's concept of "domain" becomes a valuable addition. Fiedler's model permits predictions of leadership effectiveness in interacting groups within certain situations. Fielder states it is easier to change elements of the situation than to change leader personality. A point to consider using Fielder's model is, who is to change the elements? A leader, of course.

Another approach in the study of the influence of the situation (S/L,F,S) has been in some type of "analysis" of the situation: This is similar to the systmatic, diagnostic approach found in operation analysis (OA) and management information systems (MIS).

Necessarily wide in scope, the present effort at gathering data, knowledge, and information about organizational situation(s) is rudimentary at best. In many ways it parallels the early efforts regarding search on leaders and followers. Presently, the chief value is recognizing that there are obvious situation influences affecting the leader and followers. While basic research and data gathering is being conducted, leaders and followers can expend energy on that segment of their task where it is appropriate within the situation. Even when under extreme situational influences there is always action possible in other areas of the Matrix.

In using situationally appropriate leader behavior, the leader bases his behavior upon his perception of the behavior of follower within the situation. Situationally, one has the Leadership Influence Matrix and techniques such as the Situational Leadership Influence Model, (Moore, 1977), *job engineering* from Fiedler (1977), decision making from Vroom (1972), Gestalt Techniques from Perls (1951), competency models, etc., upon which to base leader behavior.

The computer with its design capability of handling many variables permits extensive research and analysis of individual leader and follower behaviors. Television technology makes possible the direct feedback and concepts and constructs of leader, follower, and situational influence to be operationalized in behavioral terms.

Organizations may be viewed as a series and linking of followers and leaders in various arrays and positions. By systematically establishing the leader and followers in each specific sub-unit, a determination of follower and situation influence can be made in order that appropriate leader behavior may be taken. The sub-units can then be reassembled and the system viewed as a whole again.

The ability to diagnose organizational behavior in terms of the Leadership Influence Matrix (LIM) provides leaders and followers with an appropriate basis for personal behavior within systems.

All segments of personal and group life require leadership, and the more structured or institutionalized the setting, such as the church, the university, or the military, the more a systematic approach such as Leadership Influence Matrix and the Situational Leadership Influence Model assists in increased effectiveness or constructive change. Similarly, the unstructured organizations and instituitons, with their low levels of follower influence development, require initial structuring with leader behavior based upon the best data. LIM and SLIM provide to those who would be leaders or those who would teach others to be leaders and followers; a dynamic, pro-active, functional basis for leader behavior in today's and tommorrow's organizations.

REFERENCES

Fiedler, F.E., et al, *Improving Leadership Effectiveness*. New York: John Wiley and Sons, 1976.

Fiedler, F.E. *Leadership*. New York: General Learning Corp., 1971.

Forrester, J.W. *World Dynamics*. Cambridge, Mass.: Wright Allen Press, 1971.

Harris, T.A. *I'm OK—You're OK: A Practical Guide to Transactional Analysis*. New York: Harper and Row, 1967.

Hersey, P. and Blanchard, K.H., *Management of Organizational Behavior*. New Jersey: Prentice Hall, 1976.

May, R., *Power and Innocence*. New York: 1971.

Moore, Loren I., "The Situational Leadership Influence Model", Unpublished Project Study, University of Carolina: 1977.

Perls, F., Hefferline, R. and Goodman, P. *Gestalt Therapy*. New York: Julian Press, 1951.

Thompson, J.D. *Organizations in Action*. New Yook: McGraw-Hill, Inc., 1967.

Vroom, V. and Yetton, P. Toward a Normative Model of Leadership Style. *Organizational Effectiveness Research Program Technical Report* #1. Carnegie-Mellon University, Pittsburgh, Pennsylvania, 1972.

Weinstein, G. and Fantini, M. (Eds.). "Toward Humanistic Education: A Curriculum of Affect." A Ford Foundation Report. New York: Prager, 1970.

PERSONALITY INTEGRATION AND LEADERSHIP POTENTIAL: A COMPARISON OF TWO EVALUATION PROCEDURES

Major Bruce T. Caine, PhD
United States Military Academy

Individuals identified by Duncan's Personality Integration Reputation Test (PIRT), which assesses members of long-term groups for optimal functioning in the total social environment, will be compared to individuals identified as high in leadership potential by the USMA Leadership Assessment System (LAS). The latter system is the latest in a series of formalized evaluative procedures by which a cadet's primary reference group, the cadet company, report on the cadet's suitability for military leadership. It is proposed that the Academy's emphasis on leadership and its "whole person" developmental concept foster an environment in which the optimally functioning individual should also be seen as high in leadership potential. A series of instruments judging cognitive complexity, locus of control, openness to social experience, perception of social feedback, and self esteem will be administered to those cadets identified as integrated by PIRT procedure, those top quarter in LAS and a randomly selected control group. Between-group comparisons will be made. Those nominated by PIRT or LAS but not both will be compared with those jointly identified and with the control group on the measures listed above.

The development of future officers for the United States Army is the fundamental purpose of the United States Military Academy. The Academy provides a four year program that encourages total growth--physical, emotional, social and intellectual. Central to this process is the acquisition and development of the resources required for effective military leadership. Our product is hopefully an individual whose behavior reflects the knowledge, skills, beliefs and values of one who both the military community and our society at large would label a "leader." Fostering and assessing this developmental process is a principal activity of a majority of the Academy's staff and faculty.

To accomplish the leadership assessment portion of its developmental program, the Academy has employed a series of evaluative procedures, the most recent of which is the Cadet Leadership Assessment System (CLAS). The purpose of CLAS is to provide a complete and objective apprasial of the performance and potential of each cadet in the area of military leadership. Initiated in the Fall of 1977, CLAS serves four functions: counseling, recognition, practice in evaluating leadership behavior and elimination. The Cadet Leadership Assessment System relys on the cadet's primary reference group to determine suitability for military leadership. The evaluative criteria (see Figure One) employed reflect the Academy's "whole person" developmental concept and specify the ideals of this unique and demanding environment. These criteria represent the institution's definition of an "optimally functioning individual" when combined with excellence in academics and proficiency in the "social graces."

Researchers and theorists in many areas of the Behavioral Sciences have long been concerned with the assessment of the "optimally functioning individual." In the area of personality and adjustment, Jung's concept of self-actualization (later adapted by Rogers and Maslow) stressed the development and realization of all aspects of an individual into harmonious balance. In the applied area of organizational behavior and human performance, management training programs stress the integration and balance of technical, administrative and interpersonal skills. Fostering flexibility, versatility and adaptability within the climate of the organization, as well as the internalization of values (Selznick, 1957), are the objectives of many training programs.

In 1959, Seeman proposed a theoretical framework within which research in the area of optimal adjustment and effective behavior could be conducted. Duncan (1966) expanded on Seeman's framework using the concept of personality integration which is defined in terms of the quality of interaction among the human organism's various behavioral subsystems (interpersonal, cognitive, physiological, emotional/evaluative). Employing the Personality Integration Reputation Test (PIRT), a peer nomination procedure, Duncan found that the "integrated person" differs in measurable ways from his "average" peer. For example, the integrated individual was found to have a positive self-concept, to perceive himself as responsible for what happens to him, to have a wide range of interests and activities, to be intellectually efficient, to be self-directed in matters of values and beliefs, and to be positively regarded by peers. In addition, Thomas and Seeman (1971) found highly integrated persons to be cognitively complex, positively oriented towards, and less threatened by, their social environment.

Duncan's objective in preparing and testing the PIRT was to facilitate identification of those individuals in established groups seen by their peers as "having it all together", so that research on these optimally functioning persons could be conducted. Similar goals have stimulated social and organizational psychologists to conduct considerable research on those individuals identified by their peers, subordinates or superiors as leaders.

Ralph Stogdill's extensive review of this research leads him to conclude "that to a very large extent our conceptions of characteristics of leadership are culturally determined" (1974, p. 82). Within our society, leaders appear to be characterized by a strong drive for responsibility and task completion, vigor and persistence in pursuit of goals, initiative, self-confidence and a sense of personal identity, willingness to accept consequences for decisions and actions, interpersonal sensitivity, and similar cultural positives (Stogdill, 1974). Stogdill notes that "the characteristics, considered singly, hold little diagnostic or predictive significance. In combination, it would appear that they interact to generate personality dynamics advantageous to the person seeking the responsibilities of leadership" (1974, p. 82).

The parallels between the interactive complex of culturally dependent leadership characteristics cited by Stogdill and the group dependent characteristics of the integrated person as described by Seeman and Duncan are striking. For example, Seeman (1959) suggests that the integrated person makes effective use of a maximum amount of information. Shaw (Shaw and Penrod, 1962: Shaw, 1963) found efficient use of information and resistance to "information overload" to be significantly related to attributions of leadership ability in decesion making groups. Management assessment centers use devices such as "in-box exercises" to evaluate the capacity to assimilate and use quantities of information in decision making (e.g., Lopez, 1966; Slevin, 1972). Learning Fourth Class "poop" is a traditional Military Academy technique for developing information processing efficiency in new cadets.

However, previous research on highly integrated persons (Using fraternities and similar social groups) has shown that many but not all were "socially prominent" and "high achievers" in their social milieu (Duncan, 1966). The integrated person appears to be one who is able to approximate the ideals of the group to which he or she belongs. These ideals may or may not include the attaining of formal

FIGURE ONE

PIRT DIMENSIONS	CLAS DIMENSIONS
Ability to express feeling without hurting feelings of others	Work with other effectively Cooperate
Understand oneself best, aware of own strengths and shortcomings	Lead, command a group Accept authority and comply with orders.
Keep an open mind and not jump to conclusions	Sustain a high degree of loyalty while maintaining independence of thought
Best able to deal with everyday tensions and anxieties.	Organize and manage detailed assignments effectively; maintain high state of discipline
Capable of forming deep and profound relationships with others; concerned with other people.	Sustain a high degree of loyalty maintain esprit de corps and cooperation.
Most successful in all phases of life: social, personal, educational.	Participate in company and Corps activities, projects or sports. (Be academically proficient)

offices of leadership or similar highly visible social positions (e.g., varsity athletes, top scholars). Rather than "starring" on a single dimension, it appears the integrated person successfully combines all the salient dimensions of life to a degree admired and desired by others. From the available evidence, it is likely that the group's climate, interests, goals,and norms set the standards by which its members judge personal integration. Therefore, in an environment that stresses leadership development, it is reasonable to assume that the criteria for evaluating the integrated individual would be influenced by this emphasis.

A reasonable hypothesis to be derived form this evidence is that in an environment where leadership development is stressed, those seen as leaders (or as possessing leadership potential) in their established groups are also likely to be identified as integrated or optimally functioning individuals. A comparison of the PIRT and CLAS dimensions is presented in Figure One.

Method

Subjects

All members of two cadet companies selected at random from the Corps of Cadets will serve as subjects for the study. Male and female cadets of all four classes live in the same barracks area in close proximity. They eat meals and participate in sports by company. Their degree of contact can be favorably compared with the long term groups (e.g., fraternities) used by Duncan (1966).

Leadership Evaluation

The nomination system used by the cadet component of the CLAS is principly a rating of organizational subordinates. As such, it parallels the officer efficiency report system and differs from the previously employed cadet Leadership Evaluation System (LES) that placed heavier emphasis on within-class ratings. CLAS ratings are made twice during the academic year. Spring 1978 ratings will be used in this study to identify members of the subject companies seen as being in the top quarter of their class.

Personality Integration

Concurrent with the CLAS ratings, subjects will be asked to complete Duncan's PIRT nomination questionnaire. Cadets will be requested to identify three members of their own class and three members from the other three classes on each of the six PIRT dimensions. This will permit comparisons with CLAS data by year group. Based on this measure, members of the subject companies seen by peers as highly integrated will be identified. A contrast group of "average" peers (those not identified by the PIRT system) will be randomly selected from each company.

The PIRT scale employed in this project includes a self-rating measure requesting the subject to estimate how frequently peers will select the subject on each of the six dimensions. This will provide an evaluation of openness to social feedback. (Americaner, 1977)

Supporting Scales

Following evaluation of both CLAS and PERT data, cadets identified as "top quarter" in leadership, as highly integrated, or both and the random contrast group will be asked to complete the following scales:

(1) Milller and Bieri's REP Grid, a semantic differential type scale used to judge complexity in construing one's social world.
(2) Ziller's Complexity of Self-Concept Instrument, a list of self-descriptive adjectives measuring another dimension of complexity.
(3) Ziller's Self-Other Orientation Inventory measures of self-esteem, openness to experience and social space.
(4) Rotter's I-E Scale, a measure of locus of control.

Procedure for Comparisons

Based on the results of the PIRT and CLAS ratings, subjects will be assigned to one of four groups: Those identified by both PIRT and CLAS will be assigned to the DUAL group. Cadets identified by PIRT but not CLAS and vice versa will be placed in separate categories. A contrast group of individuals not identified by either procedure will be randomly selected from each company. If the CLAS system is, as predicted, identifying those cadets seen by peers as highly integrated and optimally functioning individuals, then there should be a high frequency of dual identifications. Between group comparisons will be made using a one-way analysis of variance across the four groups for each of the supporting scales and the self-rating of PIRT identification to determine what factors clearly differentiate among these groups.

It is possible that other factors may have an impact on PIRT and CLAS ratings. Since this research is being conducted with existing groups, major sociometric variables such as sex, age, race, class, and room location cannot be manipulated. These factors will be investigated, as will academic performance and formal leadership position, for systematic differences.

Summary

This investigation will provide an external validity check on the Cadet Leadership Assessment System and may establish a theoretical and empirical link between the concepts of military leadership and optimal personal functioning. It will also serve to enhance our understanding of the cadet perspective on effective behavior in the West Point environment.

REFERENCES

Amerikaner, M. Optimal personality functioning: A general systems theory approach. Research Proposal, University of Florida, 1977.

Duncan, C.B. A reputation test of personality integration. *Journal of Personality and Social Psychology*, 1966, 3 (5), 516-524.

Hearn, C.B. & Seeman, J. Personality integration and perception of interpersonal realtionships. *Journal of Personality and Social Psychology*, 1971, 18 (2), 138-143.

Kelley, H.H. Two functions of reference groups. In G.E. Swanson, T.M. Newcomb, and E.L. Hartley (Eds.), *Readings in Social Psychology*. New York: Holt, Rinehart, and Winston, 1952, pp. 410-414.

Lopez, F.M. Jr. Evaluating executive decision making: The in basket technique. *AMA Research Study* 75, New York: American Management Association, 1966.

Miller, H. and Bieri, J. Cognitive Complexity as a function of the significance of the stimulus object being judged, *Psychological Report*, 1965, 16 (3),1203-1204.

Rotter, J.B. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, 1966, 80 (1), 1-28.

Rotter, J.B. & Mulry, R.C. Internal versus external control of reinforcement and decision time. *Journal of Personality & Social Psychology*, 1965, 2 (4), 598-604.

Seeman, J. Toward a concept of personality integration. *American Psychologist*, 1959, 14, 633-637.

Seeman, J. Teacher judgements of high and low adjustment. *Journal of Educational Research*, 1963, 57, 213-216.

Selznick, Philip, *Leadership in Administration*. New York: Harper & Row, 1957.

Shaw, M.E. Some effects of varying amounts of information exclusively possessed by a group member upon his behavior in the group. *Journal of General Psychology*, 1963, 68, 71-79.

Shaw, M.E. and Penrod, W.T. Does more information available to a group always improve group performance? *Sociometry*, 1962, 25, 372-390.

Slevin, D.P. The Assessment Center: Breakthrough in management appraisal and development. Personnel Journal, 1972, 51, 255-261.

Stogdill, R.M. *Handbook of Leadership, A Survey of Theory and Research*. New York: The Free Press, 1974.

The Operation and Administration of the Cadet Leadership Assessment System. Draft Regulation, USCC, undated.

Ziller, R.C. Self-Other orientation theory and measures. Manuscript, University of Florida, 1973.

Ziller, R.C. *The Social Self*. New York: Pergaman Press, 1973.

THE NAVY LEADERSHIP AND MANAGEMENT TRAINING PROGRAM: A COMPETENCY-BASED APPROACH

Paper presented at the sixth Symposium on Psychology
in the Department of Defense, United States Air Force Academy,
Colorado Springs, Colorado
20 April 1978

Lyle M. Spencer, Jr., Ph.D.
McBer and Company
137 Newbury Street
Boston, Massachusetts 02116

Introduction

This paper will briefly summarize the research underlying the Navy's new Leadership and Management Training Programs.

In 1975, under the leadership of CDR Dana French, the Navy's Bureau of Personnel (BUPERS) initiated a study of the Navy's leadership and management training efforts. This study (CNET,1975) produced a number of interesting findings.

First, it was found that a majority of naval commissioned and non-commissioned officers did not receive any leadership and management training at key ascension points in their careers.

Figure 1 shows the key ascention points for commissioned officers—source school conversion training when people enter the service, then division officer, department head, staff, XO, CO, and Flag positions, and shows the number of empty cells, points at which officers do not now receive training.

Figure 2 shows the comparable ascension points for enlisted personnel—recruit (boot camp training), petty officer, leading petty officer, leading chief petty officer, and master chief petty officer (E9s in the position of senior enlisted person in command)—and shows the empty cells, points at which noncommissioned officers do not receive training.

Even where leadership training is shown to exist, it is often very limited (for example, as few as four hours in a two-month technical school). The conclusion was that most Navy leaders do not receive adequate training in leadership and management at the key points in their careers when they must assume increased management responsibility.

Second, BUPERS identified some 157 different leadership courses in various commands, plus numerous "unofficial" courses sponsored by commands "out of hide" (i.e., out of their own funds), one unobtrusive indicator of the need for management training felt by Navy personnel at all levels.

Third, an examination of the content of existing courses found:

- There was no standard curricula or consensus about what knowledge or skills were needed to perform effectively as a Navy leader. Courses consisted of a potpourri of Navy tradition, rules, and regulations; civilian academic management theories (e.g., Maslow, Hertzberg & McGreggor); and a few offerings from the behavioral sciences—organizational behavior, communications, transactional analysis, and the like.

- Courses lacked any foundation in empirical research (i.e., there was no data to show that anything taught in these courses had, in fact, any relation to effective leadership and management in Navy billets).

- Courses were 90 percent cognitive and 10 percent experiential—i.e., consisted primarily in talking at people *about* leadership and management in formal classroom lecture formats rather than teaching them actual leadership and management skills.

Most Navy participants reported that they liked these courses, but that they experienced difficulty relating what they had learned to their actual jobs.

The Research Effort to Identify Navy Leadership Competencies

To correct these limitations in existing courses, BUPERS decided to develop a new leadership and management program based on empirically derived competencies: those knowledge, skill, and motivation variables which could actually be shown to predict effective performance in Navy leadership billets.

The research method used to identify these competencies was a job competency assessment procedure developed by Professor David C. McClelland at Harvard University (McClelland, 1975). This procedure consists of three steps (Figure 3a):

1. identification of a criterion sample of exemplary (top five percent), superior, and average billet incumbents at each ascension point from the three Navy communities (air, surface, and submarine commands) in the Atlantic and Pacific Fleets.

2. behavioral event interviews with sample subjects and

Lyle M. Spencer, Jr., Ph.D. is Vice President of McBer and Company, a Boston consulting firm which developed the job competency assessment approach. Additional copies of this paper and the technical reports on which it is based can be obtained from McBer and Company, 137 Newbury Street, Boston, MA 02116.

OFFICER COMMUNITY/KEY BILLETS SUMMARY

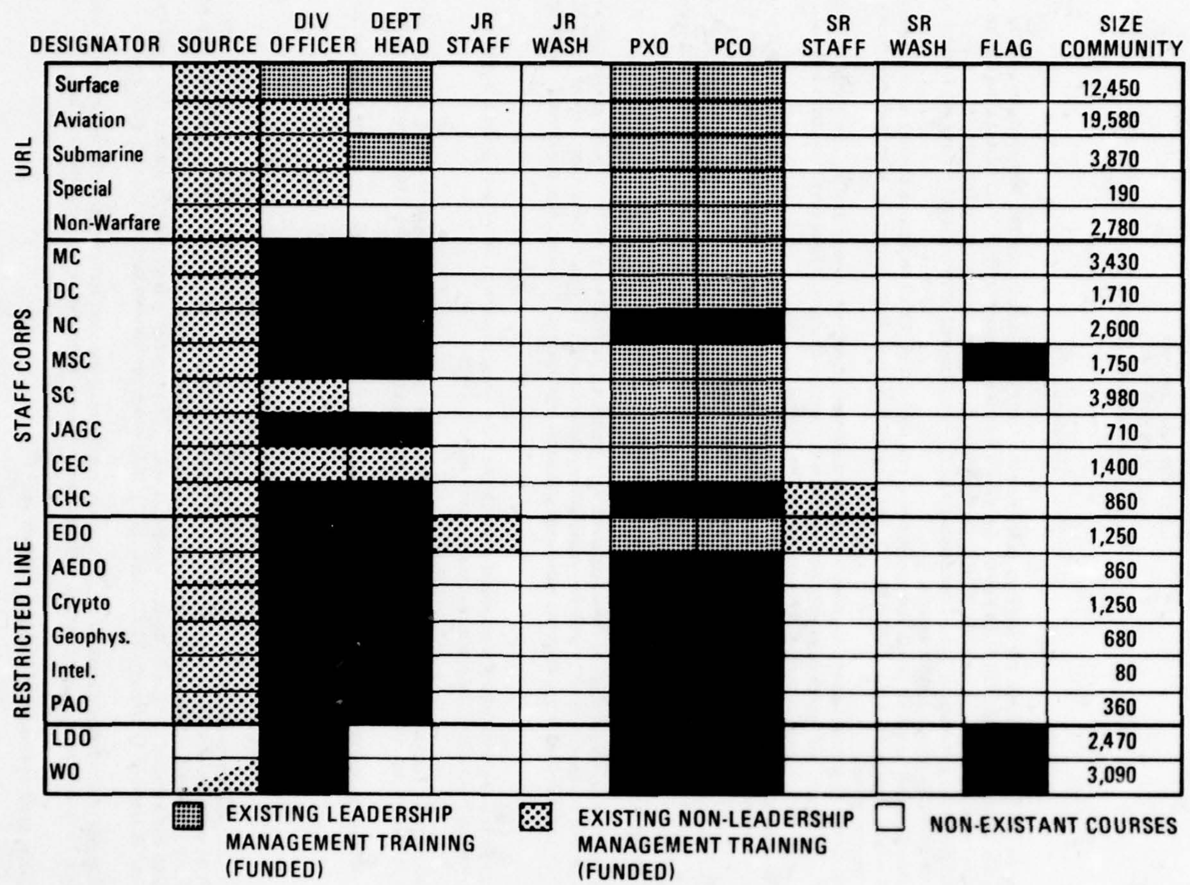


FIGURE 1

ENLISTED OCCUPATIONAL FIELDS/KEY BILLETS SUMMARY

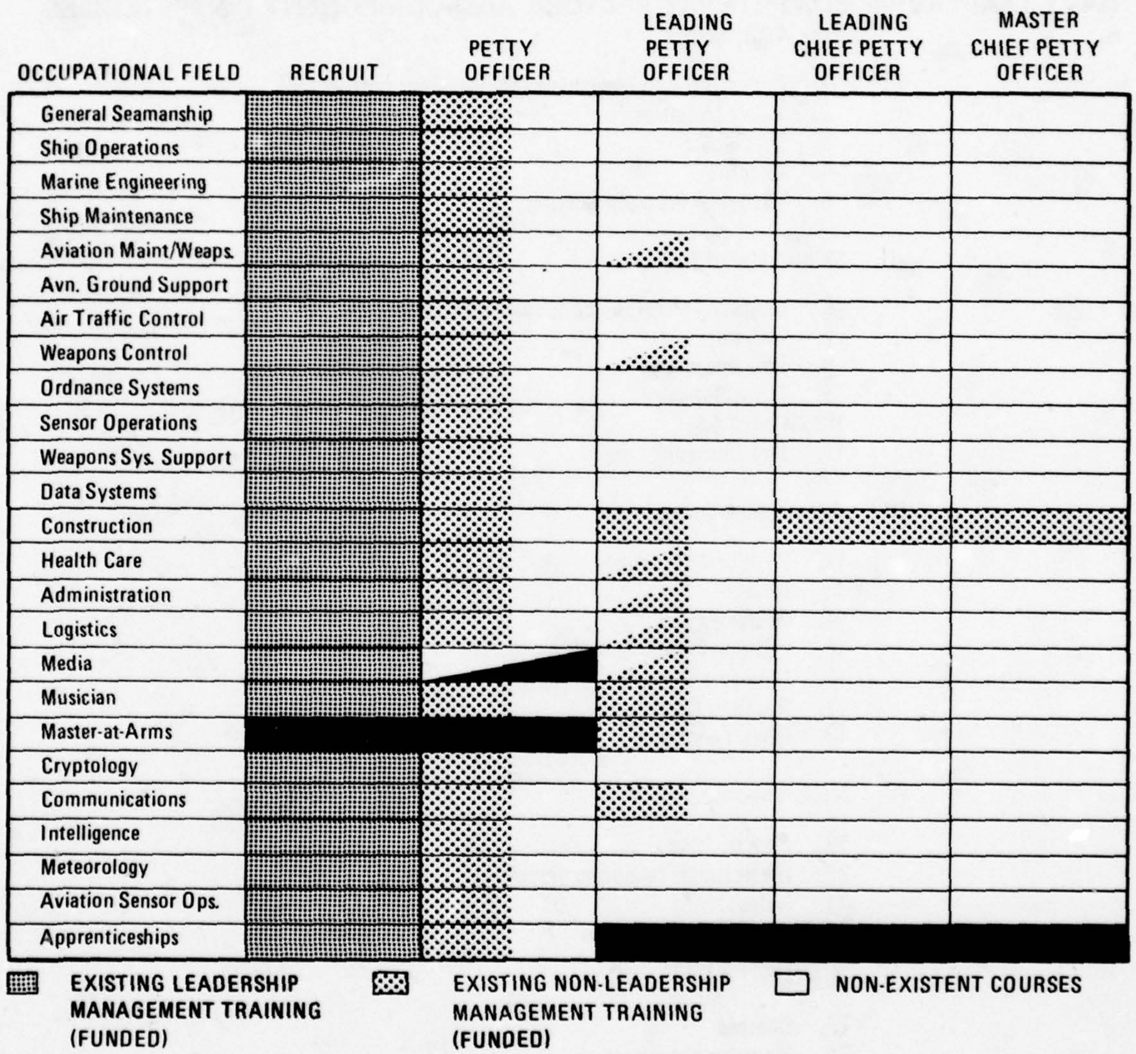


FIGURE 2

FIGURE 3a
STEPS IN THE JOB COMPETENCY ASSESSMENT PROCESS

1. Criterion Sample: superior vs. average job incumbents
2. Behavioral Event Interviews
3. Thematic Analysis

- Causal personality variables (e.g., motives) as well as knowledge and behavioral skills
- Empirical coding

FIGURE 3b
NAVY LEADERSHIP COMPETENCY FACTORS AND COMPONENT COMPETENCIES

I. Task Achievement

1. Concern for Achievement (achievement motivation)
2. Takes Initiative
3. Sets Goals
4. Coaches
5. Technical Problem Solving

II. Skillful Use of Influence

6. Concern for Influence (power motivation)
7. Influences
8. Conceptualizes
9. Team Builds
10. Rewards
11. Self Control

III. Management Control

12. Plans and Organizes
13. Directs
14. Delegates
15. Matches People and Jobs
16. Monitors Results
17. Resolves Conflicts
18. Gives Feedback

IV. Advises and Counsels

19. Listens
20. Understands (accurate empathy)
21. Helps
22. Positive Expectations

V. Coercion

23. Coerces
24. Negative Expectations
25. Disciplines
26. Acts Impulsively
27. Fails to Resolve Conflicts

3. thematic analysis of interview protocols.

The behavioral event interview process is similar to Flanagan's (1954) familiar critical incident technique, with three refinements. First, it assesses specific individuals in known criterion groups, rather than using groups to generate incidents typical of a given job. Second, it identifies underlying "causal" traits, such as motivation and cognitive skills, which predict job performance, while the critical incident technique records only behavior. Interviewers use a structured inquiry procedure similar to that used in administering a verbal thematic apperception test (TAT). These probes elicit a respondent's motive, cognitive style in problem-solving, habitual evaluations of others, and other personality variables.

Third, individual characteristics in interview data are empirically coded, providing quantitative scores that can be used in statistical tests of their power to predict subjects' criterion ratings.

Approximately 200 respondents were interviewed, yielding about 900 scorable behavioral events. Thematic analysis of a subsample of these protocols identified 29 reliably identifiable competency characteristics. All events described by rated officers in the Atlantic and Pacific Fleet samples were then scored "blind" by trained scorers who had achieved a mean interrater reliability of r = .86. Competency variable scores for all subjects were then factor analylzed, yielding five competency factors (Spencer, 1977; Klemp, Munger & Spencer, 1977)[4].

The first competency factor was *task achievement* (Figure 3b.1). Superior officers reported more incidents in which they expressed concern for achievement, set specific goals, took initiative to solve technical problems, and coached others to improve their performance.

The following incidents illustrate the competencies involved in the Task Achievement factor.

> (Aircraft squadron CO): "When I took over the squadron, the first thing I did was get all the flight records going back six months. We weren't doing nearly as well as the other squadrons, and I told the wardroom that I said I wanted us to be the best, and that we should increase our flight hours every week until we were. Mostly it was a matter of showing people how to plan, to count the days, figure out what to accomplish each day, personnel, materials, who goes where, when, what needs to be done, when, and who does it, down to the lowest airman. I got all the LCDRs together, showed them how to look at all elements that were going to occur in the future, how to develop a plan of attack. . . Six months later, we've got the most hours flown in the group."

> (EN3): "This one crane kept going down. I found it was a nut that would shake loose, dropping the whole subassembly. I wanted to fix it right for once--I figured a way to put a lock nut on it and a cotter pin behind that, so it couldn't shake off. It's worked real well since, hasn't been down. That made me feel real good."

In the first case, the CO expressed a concern with improving performance, setting goals, and achieving his objective by coaching his subordinates. In the second case, achievement concern led the EN3 to take the initiative to solve a technical problem.

The second competency factor was *skillful use of influence* (Figure 3b.2). Superior officers were more concerned with influence (i.e., higher power in motivation), and used more influence methods--persuasion, explanation, inspiration, rewards--to accomplish their objectives and motivate subordinates to work as a team.

The following incident illustrates the competencies involved in the Skillful Use of Influence factor:

> (RM7): "I went in there, junior to everyone else, and I knew I had a better way. These people'd been doing it the same way for 15 years--it was their baby and they had real ownership in it. I wanted subtly to convince them my way was best. I did this by proposing new ways that were almost complete, but which they could make a real contribution to--so they owned it, too. I kept my cool and let them get lots of wins off me on little stuff so they always felt they were on top--though I got the big ones I wanted through."

In this example, the officer expressed concern with influence, controlled his emotions, and effectively influenced people and events in the direction he wants, while leaving his team members feeling rewarded and strong.

The third factor is *management control* (figure 3b.3). Superior officers were more likely to report incidents in which they planned and organized tasks, provided clear direction, matched people to jobs, delegated responsibility, monitored or followed up to be sure jobs got done, and provided subordinates with feedback on their performance.

The following incident illustrates this factor:

> (CO, submarine): "I started planning for the inspection six months ahead of time. If you don't do that, you're always fighting fires and you never get on top of it. Crisis management I don't like. I laid it all out ahead of time (shows interviewer elaborate PERT chart). I make sure every guy knows his job. (Interviewer: "How?"). . . I tell 'em. . . each gets his chunk of the plan. . . (shows interviewer thick folder with "to do" lists for each department) and knows where he fits with the others. I let them run with it, but haul them in every week to see where they're at, what's going okay and what ain't. If two guys don't agree on something, they have a little chat with the XO or me. I don't let things slip. I've got a tickler system (shows interviewer three-inch stack of NCR forms referenced by due dates)."

---

[4]Technical reports which include the description of the 27 competencies, the scoring manual for coding them in interview protocols, methods used to establish interrater reliability, and statistical analyses of data, may be obtained from McBer and Company, 137 Newbury Street, Boston, MA 02116.

Superior officers reported planning ahead and feeling in control of events. Average officers expressed feelings of always being behind, overwhelmed by external system demands, and powerless to avoid continued crisis management.

The fourth competency factor was *advising and counseling* (Figure 3b.4). Somewhat to our surprise, we found that Navy leaders spent considerable time acting as counselors, advising subordinates on performance, disciplinary, career planning, and personal problems such as drug, alcohol, financial, and family difficulties.

Superior officers reported incidents in which they identified subordinate problems, initiated contact to talk about these problems, or were perceived as approachable. They described their ability to hear what people were tyring to say (accurate empathy or insight into others' needs, motives, and hidden agendas), and acted to help subordinates with personal problems by "fighting for their people" or making appropriate referrals to other sources of help. These officers also made significantly more statements of belief in subordinates' basic worth and ability to perform, and more use of positive expectations to inspire imporved subordinate performance.

> The following critical incident illustrates this factor:
>
> (BM2): "I had this one seaman, a real good kid. Suddenly his performance started slipping. I'd see him wander off by himself during breaks and just stare out to sea. I went up to him and said 'I don't want to mess with your private life, but have you got anything you'd like to talk about, sort of get off your chest?' He said he was going to kill himself. He'd gotten a letter from his wife saying she was leaving him, living with some other guy. I talked with him a lot after that, over the next few months—mostly just listened. He got divorced, but he didn't do himself no harm. He's okay now. He's one of my best men."

In this case, the petty officer noticed his subordinate's distress, made contact in an understanding way, provided counseling assistance, and expressed consistent positive regard for the individual throughout the process.

The final factor is *coercion* (3b.5) use of threats or rank to motivate others to act, expression of negative expectations of subordinates, disciplinary actions, and resolution of conflicts by force in "win-lose" confrontations. Interestingly, this factor did not differentiate superior from average leaders. Superior officers actually reported more incidents in which they disciplined subordinates. Perhaps because they are more concerned with standards and management control, these officers gave more feedback, both positive (rewards) and negative (disciplinary).

Coercion appears to be effective when it is required to maintain standards or order essential for mission performance—for example, being very tough on drug abuse on nuclear submarines, as illustrated in the following incident.

> (XO), submarine): "I know there are people doing drugs on board. I tell every new man who comes on board, 'Drugs don't go down on this ship—and if you mess with that stuff, I'll catch you and bust you.' You just can't have that kind of nonsense on a nuclear sub. I caught two last week and it was two and a half days from the time they were caught to the time they were *gone*. And if you know the UCMJ red tape involved, that's fast."

Coercion tends to be ineffective when used solely to maintain a superior's power position or to demand conformity on minor issues not related to mission accomplishment (characterized by respondents as "nit-picking" or "harassing people about Mickey Mouse stuff"). In dealing with subordinates' initial problems, errors, or infractions, coercion appears to be less effective than restraint followed by counseling or coaching. Superior officers tend to coerce subordinates only as a last resort, when the other leadership methods prove ineffective, whereas average officers are more likely to use coercion as a first or only resort. This distinction was illustrated by two officers interviewed the same day in San Diego. Each had received seamen who had been in considerable trouble in their former units. The first officer reported:

> "I told that kid that I didn't care about his past, that he started at day one with me, but that if he messed up, I'd break him personally. I got on him about his appearance. He was a habitual pig-pen. One day I saw him holding a slack line on a capstan. I told him to take up the slack. He just stood there. I said, 'All right, you've had it.' I wrote him up on two counts of insubordination and failing to obey an order. He filed a complaint against me charging me with harassment. Then he went UA. Two days later he was found dead in a bathtub of an overdose of drugs. Maybe I helped cause it, but what can you do with someone like that?"

The second officer said:

> "I get transferred a lot of so-called 'deadheads,' but they aren't really, 'cause every man can do some job well, and I can prove it. I told this one kid I didn't care about his past, he started clean here and I'd help him every way I could. I gave him simple jobs I knew he could succeed at to give him some wins. I tried to think of things to do to build his self-confidence. Everyone else had a tool kit, so I gave him one. Damn if he didn't break down and cry. He said no one had ever trusted him enough to give him something like that before. It took a while, but he's become an okay mechanic and he's not been in any trouble since he's been here."

The first officer, rated average, expressed negative expectations and used coercion resulting in an undesirable outcome. The second officer, rated exemplary, expressed positive expectations, and used a coaching approach which effectively empowered his subordinate and resulted in the latter becoming an adequate performer.

A final significant finding was that superior officers have a wider repertoire of leadership competencies than average officers: They report use of more and different leadership competencies, suggesting greater flexibility in responding to situational contingencies. Superior leaders differ not in coercing *less*, but in coaching, influencing, managing, and counseling *more*.

The competency factors were tested in a double cross validation design in which the discriminant function analysis coefficients predicting superior performance in one Navy fleet were used to predict ratings in the opposite fleet. Stepwise discriminant analysis yielded standardized coefficients for all competency factors except Coercion, which, as noted, did not distinguish superior from average officers.

The PACFLT discriminant function correctly classified 41 of the 51 rated officers (80.4 percent) in the PACFLT sample and 55 of the 78 officers in the LANTFLT sample (70.6 percent). The LANTFLT discriminant function correctly classified 58 of the 78 rated officers (74.4 percent) in the LANTFLT sample and 36 of the 51 rated officers (70.6 percent) in the PACFLT sample. The average of the *phi* coefficients obtained in each cross-predicted sample is .46, significant at the $p < .001$ level and a reasonable estimate of the predictive power of the four-factor competency model when applied to an independent sample. No significant rating or competency differences were found for superior officers in the Atlantic and Pacific Fleets or among the air, surface, or submarine communities. These data suggest that the rating standards were relatively unbiased and that competencies identified are applicable to the Navy as a whole. These findings provide an empirical basis and method for the design of Navy leadership and management assessment tests and training programs.

Design of Navy Leadership and Management Education and Training (LMET) Programs

I will conclude by briefly describing how these research data are being used in the design of competency-based leadership and management courses. Each ten day course includes four parts, as illustrated in Figure 4a.

The first part of the course is *self-assessment*. Trainees receive feedback on their performance on each of the competency factors predicting superior performance. This feedback, which in effect tells them "this is how the superstars in your billet perform on these dimensions, and this is how you perform," provides each trainee with very specific directions or goals for change. Descriptions of what superior performers actually do to get their results—and the implication that trainees can learn skills—provides a potent role model and motivation for participation in the course.

The second part of the course consists of *intensive modules on each competency*. Each module includes four elements (Figure 4b):

1. *Recognition*. Trainees are presented with a difficult case or simulation problem designed to produce a "shock of recognition," in which participants realize that they actually encounter problems calling for use of the competency in question in their real jobs and that they often have trouble dealing with these problems. (For example, in a module on influence, trainees may be asked to motivate a subordinate who panics and refuses to perform in a crisis.) These cases and simulations are drawn from actual critical incidents collected during the research phase of the project. The recognition element is designed to introduce the topic in an experiential way and create specific motivation for learning: maximum relevance with a recognition of actual versus desired skill in the competency.

2. *Explanation*. Trainees receive practical conceptual input on the competency via readings, lectures or demonstrations. For example, in a module on influence, participants would get a lecture on the importance of power motivation in leadership tasks, how to see managerial interactions in terms of influence skills, then a videotape demonstration of specific ways of dealing with the situation presented in the "recognition" case.

3. *Skill Practice*. Trainees practice using the competency skills they have been taught. For example, they discuss additional influence cases, describing how they would handle each, then actually practice giving charismatic speeches, persuading subordinates to reenlist, dealing authoritatively with crisis situations, and the like. Videotape critiques are used to give participants specific feedback on their development of skills.

FIGURE 4a
DESIGN OF TEN DAY NAVY LEADERSHIP AND MANAGEMENT COURSE

Day 1 : Self-assessment

Days 2-6: Intensive modules on each competency factor

Days 7-9: Large scale job simulation

Day 10 : Goal-setting and post-test assessment

FIGURE 4b
ELEMENTS OF EACH COMPETENCY FACTOR MODULE

1. Recognition
2. Explanation
3. Skill development
4. Practical application

4. *Job Application.* Trainees identify situations in which they will use the competency in their jobs and set goals, anticipate obstacles, and develop action plans for doing so. This element reinforces the relevance of training, translates it back into real job performance, and increases the likelihood that skills developed in training will in fact be used back on the job.

The third part of the course consits of a *large scale job application simulation.*

Because it is recognized that jobs rarely require the use of a single competency, a three-day full-scale simulation was designed to present participants with complex problems requiring the use of many of the competencies simultaneously. For example, in a course for commanding officers, participants are required to take over a command facing a major inspection where the previous CO has been relieved for cause. They first must deliver a "CO policy statement" (an influence task). They then must plan for the inspection, a task which requires negotiation with their superior officer and various supply units for material and resources, under time pressure and in competition with other COs who need the same limited resources (a task with technical achievement, influence, and management control elements). As they manage their subordinates' preparation for the inspection, they encounter various obstacles: a major equipment casualty, a racial incident which threatens to embarrass the CMD and bring work to a halt, and the breakdown under stress of a key subordinate who goes AWOL on an alcoholic binge (problems requiring the use of technical achievement, management, influence, and counseling competencies in various combinations). Each problem is realistic, drawn verbatim from critical incidents of actual billet incumbents.

The final part of the course consists of *goal-setting and evaluation.* Trainees are post-tested on the competencies, helped to integrate material learned, and asked to set goals for use of developed competency when they return to their jobs. Goal-setting exercises have a career and life planning emphasis, encouraging ongoing self-assessment and planning for personal development. Post-tests provide data for evaluating the course.

In summary, the Navy's competency-based leadership and management program is designed to teach practical leadership skills which have been shown empirically to predict competent job performance: what *superior* leaders *actually do* in their real jobs.

REFERENCES

CNET. *The current status and operational characteristics of leadership and management education and training.* Pensacola, FL: Office of the Chief of Naval Education and Training, November, 1975.

McClelland, D.C. *A guide to job competence assessment.* Boston, MA: McBer and Company, 1976.

Klemp, G.O., Munger, M.T., & Spencer, L.M. *Analysis of leadership and management competencies of commissioned and noncommissioned naval officers in the Atlantic and Pacific Fleets.* Boston, MA: McBer and Company, 1977.

Spencer, L.M. *The Navy leadership and management education and training program.* Boston, MA: McBer and Company, 1977.

THE U.S. MARINE CORPS' LEADERSHIP EVALUATION AND ANALYSIS PROGRAM (LEAP): A SELF-DEVELOPMENT APPROACH TOWARD ORGANIZATIONAL EFFICIENCY

Thomas D. Affourtit
Interaction Research Institute

The United States Marine Corps would like to introduce a new program of leadership development . . . the Leadership Evaluation and Analysis Program, commonly known as the LEAP.

- The LEAP is designed for use by the small-unit commander to identify leadership concerns, to measure overall unit combat readiness, and to evaluate the effectiveness of the decision-making process.

- The LEAP provides company-, battery-, and squadron-level commanders with a completely decentralized leadership aid. The program is self-applied and voluntary on the leader's part. And, the results are strictly confidential to the individual command.

- Most important, through the decision-making feedback principle, the LEAP aids the leader in developing the flexibility to deal with various groups under a variety of situations and mission requirements.

The LEAP is basically an intelligence-gathering process founded on the principles of modern management methods, and utilizing behavioral science techniques. The entire program is presented in a manual that features a step-by-step procedure for administering the techniques and explicit guidelines for scoring and reviewing the results. Therefore, no outside professional assistance is necessary and no reports or formal paperwork is required.

Program materials allow the commander to systematically measure unit performance in terms of general Marine Corps standards and specific unit requirements. In addition, unit motivation is measured to determine the reasons behind various levels of performances. Since performance is the consequence of a motivational state, the LEAP attempts to measure the causes as well as the effects of unit combat readiness.

Once the causes of performance levels are identified, the leader can take the necessary action to extinguish those conditions that produce deficient behavior, and reinforce or support the conditions that promote positive performance.

The technique for measuring unit motivation is a questionnaire that functions much like a starlite scope. It enables the leader to see the hidden causes of performance that are not easily observed under normal conditions. In this way, the LEAP procedure is not unlike counterinsurgent or search-and-destroy operations. But, in this case, the enemy is indifference, negligence, discord, and prejudice. These are some of the reasons that separate effective, combat-ready units from ineffective units.

The application of the LEAP is very simple, requiring only a few minutes of the commander's time to request application, to review the results, and then to make appropriate counteractive decisions.

First, the commander has the unit clerk record various performance statistics on the Leadership Analysis Form, a behavioral measuring technique and an essential part of the program. This part is easily accomplished since all the information needed is available from unit records. It takes about one hour to accomplish.

Second, the CO directs a responsible subordinate leader to administer the Interaction Inventory, the motivational questionnaire, to the entire command to measure the motivational level of several relevant areas of Marine Corps concern. It takes from 15 to 35 minutes for troops to complete the anonymous questionnaire. Questionnaire results may be either scored manually or computer processed.

Next, the CO reviews the results of the questionnaire according to a scheme outlined in the manual, or according to his or her own interest. These results serve as benchmarks for the command that are used to judge progress in critical leadership areas.

Let's take a closer look at how the LEAP works.

The CO calculates his or her own performance profile, a procedure that is similar to developing the readiness indicators used at some higher level commands, except that unit leaders select their own areas of importance, and unit measures are taken over a designated period of time.

The command motivational profile is then reviewed to determine areas of strength and weakness in terms of troop perceptions. Several general areas are measured, such as command efficiency and cohesion, that indicate a level of unit preparedness. And, conditions of discrimination, justice, and inter-group climate provide an assessment of overall command equality.

Like an aerial reconnaissance photo, this information can be amplified to expose specific issues and conditions that make up each scale in the command motivational profile. Again, strengths and weaknesses are noted as peaks or depressions along a scale graph. Tactically speaking, this information gives the coordinates of the enemy's position. These data also allow the leader to establish mission objectives or goals in order to destroy the negative condition and improve unit military status.

Validity studies conducted on Marine Corps field units show that as these scale scores move toward the right, indicating more positive motivational levels, unauthorized absences within a command significantly decrease, and first-term reenlistment rates significantly increase.

The motivational information can also be analyzed from the standpoint of any number of groups within the command, such as senior versus subordinate Marines, mirority versus majority troops, or by educational level, marital status, or career intention groups, just to name a few.

It has been found that where disparity exists between groups distinguihed by rank, such as E-5 and below versus E-6 and above, absenteeism will again be high and reenlistments will be low. But, when the scores of the two groups move closer together, indicating that both groups judge the command as functioning at the same level, unauthorized absences again decrease significantly and reenlistments increase significantly. The same outcome was discovered for differences in perception between minority and majority member troops.

In review, there are just four primary steps in the LEAP procudure:

1. Identify leadership problem areas or locate enemy positions using the LEAP tactical sensing and reconnaissance devices (analyze).
2. Establish leadership goals and management objectives (plan).

3. Take corrective action (attack).
4. Evaluate results in terms of performance and or motivational outcomes (evaluate).

In this same manner any leadership decision, program of training or policy designed to improve performance can be evaluated. Such decision-making evaluation assists Marines in producing viable solutions to some of the contemporary porblems that all Marines face.

In support of this solution-oriented approach, an information processing data bank is planned. The LEAP Network Monitor System is a process that will enable Marines to share solutions to common problems encountered in the field. This system will function on a voluntary and anonymous basis as a data input and inquiry bank for Marines using the LEAP. Information input, recommendations, and solutions discovered can be analyzed, and data feedback can be presented in consideration of any number of influencing conditions, such as unit composition, mission, unit status, location, or effective strength. Such results, based on actual conditons, will also be beneficial in training new leaders to make the most appropriate and effective decisions prior to command assignment.

In essence, the LEAP supports one of the most basic and time-proven leadership principles: know your personnel! It doesn't matter whether your leadership style is authoritarian, participatory, or charismatic; you can benefit form a better understanding of unit personnel and more systematic knowledge as to how members react under various situations.

In summary, the LEAP is a voluntary aid for company-level commanders. And, the payoff for just a little time and effort invested may be:

- Greater unit harmony and morale
- Performance improvement
- Less disciplinary effort expended
- Increased control and influence over troops
- Leadership skill development

IS THE EYE SMART OR THE BRAIN FORGIVING?

Stanley N. Roscoe and Russell A. Benel
University of Illinois at Urbana-Champaign

Somehow we see things clearly, or at least are unaware that they are unclear, when in fact the images on our retinas are badly out of focus as determined by an infrared optometer. Although the average accommodation distance may be far from the distance of the viewed object, the more or less rhythmic fluctuations in accommodation distance are sufficiently large to bring an out-of-focus object into focus momentarily every so often, at which times it may be sampled by the brain. Furthermore, if objects are brought into focus regardless of accommodation distance by the use of a small artificial pupil, the accommodative mechanism will quickly lapse toward its intermediate resting position and then, after a minute or two for many subjects, will embark upon a series of extreme fluctuations, as if searching for an out-of-focus image to back away from. These and other incidental experimental observations are now subjects of an ongoing systematic investigation.

BACKGROUND

During the second half of the last century, students of visual sensation and perception were starting to call themselves psychologists. Many of these psychologists, particularly the Germans, had taken their formal training in physics and physiology. Quite naturally they devoted much of their early attention to physical adjustments that could be observed and measured, either directly by inspection or indirectly by introspection. Unfortunately, in their compulsion to make things tidy, they bequeathed us a legacy including some untested assumptions and a few downright misunderstandings that they had brought along from physics and pshysiology.

The Misunderstandings

One assumption that has misdirected psychologists for more than a century is the misbelief that the dark focus of the eye—its relaxed accommodation distance—is at the far point, normally taken to be "optical infinity." This long-accepted "fact" was brought into question with the discovery of the phenomenon of "empty-field myopia" experienced by pilots of high-flying airplanes (Whiteside, 1957). However, it was not until the present decade that the "intermediate distance of dark focus" was firmly established by Hershel Leibowitz and his students at Pennsylvania State University (e.g., Leibowitz & Owens, 1975) and by Robert Randle and his associates at NASA's Ames Research Center (e.g., Roscoe, Olzak & Randle, 1976). Although individual resting accommodation distances vary widely, the typical distance is at arm's length.

A closely related misconception, often an implicit assumption in experiments, is the belief that the eye reflexively accommodates reasonably accurately to the distance of an object being attended to in central vision. In fact, Hennessy and Leibowitz (1971) have shown that accommodation depends upon the distance to the peripheral surroundings as well as to the foveal target; Roscoe, Randle, and their associates (Roscoe, et al., 1976; Roscoe, 1977; Randle, Roscoe, & Petitt, in press) have shown that accommodation outward or inward is a compromise between the "pull" of the stimulus and the tendency of the eye to lapse toward its resting position; and Randle (1971) has shown that, through biofeedback conditioning, accommodation can be brought under voluntary control independent of the visual stimulus.

For accurate accommodation, two conditions are necessary: (1) adequate textural and/or perspective cues to distance and (2) the requirement to make a fine discrimination. Lacking either, the lazy eye simply doesn't bother to focus, and the forgiving brain pretends not to notice. Thus, the validity of countless experiments done in darkened rooms is limited to those conditions, and the generality of the findings to the everyday visual world is suspect. Indulging the naive "scientific" compulsion to study the "pure" effect of one variable at a time, and to hold constant or eliminate the presence of all other variables, is to deny the undeniable interactive complexity of visual processes.

For examle, many investigators have used ophthalmic lenses of varying dioptric power to "induce" accommodation to different distances, not bothering to measure the actual resulting accommodation levels. Randle, et al. (in press) recently attempted such a manipulation and did measure accommodation. The investigators were surprised by the extent of the eye's disobedience and the brain's indifference. Over a stimulus range of three diopters, the eyes of 20 pilots shifted their accommodation to real images by 1.46 D, on average, and to virtual images by only 1.27 D.

The Reduced Schematic Eye

One legacy from physiological optics that remains suspect is the central assumption upon which currently accepted models of the reduced schematic eye are based, namely: that the angle subtended by the projected retinal image of an object is proportional to the visual angle subtended by the object, *regardless of the distance to which the eye is accommodated* (Davson, 1972; Elder, 1940). Both the Law of Size Constancy and Emmert's Law regarding the projection of afterimages depend upon this assumption, and both laws break down when the eye accommodates to different distances (Holloway & Boring, 1941; Young, 1948; 1952).

The proportional retinal angle assumption and an alternative hypothesis are illustrated in Figure 1. The testing of the assumption is the subject of an ongoing investigation by the authors for the Air Force Office of Scientific Research. If the assumption proves false, the way will be cleared for the explanation of unexplained findings throughout the literature of visual perception of size and distance and various bias errors in vehicle control. Available evidence from experiments by Young among others supports the rejection of the proportional-angle hypothesis and acceptance of the alternative.

THE MEASUREMENT OF ACCOMMODATION

Much of the misunderstanding and dependency upon untested assumptions concerning visual perception has resulted from the difficulty

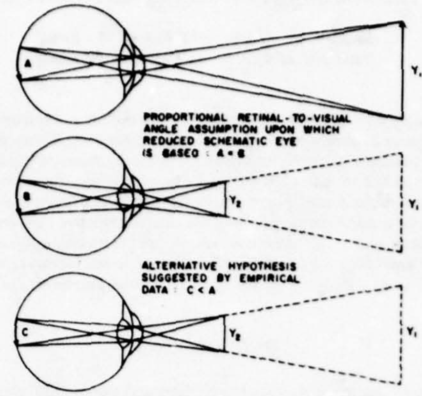


FIGURE 1. Illustration of the proportional retinal-to-visual angle assumption and the alternative hypothesis that the retinal angular projection is attenuated with increasing lens convexity.

of measuring accommodation, either overtly or covertly, while subjects are making perceptual judgments. What has changed recently is the ready availability and widespread use of devices capable of relatively accurate covert measurement of accommodation without affecting it or seriously restricting the subject's performance of perceptual tasks (Crane & Steele, 1978; Crane & Clark, 1978). Measurement techniques have been available for a much longer time, but they have not been widely used—some not at all since their original development (e.g., Wulfeck, 1952).

Overt Measurement

Leibowitz and his students and many others have made frequent and effective use of a simple and inexpensive laser optometer that requires an overt vocal or manual response by the subject (Leibowitz & Hennesy, 1975). The device has the advantage of absolute, as opposed to relative, measurement. While it does not affect the subject's accommodation, it does distract attention from a primary perceptual task. Furthermore, it does not yield continuous measurement as required for the study of the speed of accommodation or its microfluctuations.

Covert Measurement

Oculomotor adjustments that constitute part but not all of the accommodation process are changes in the curvature of the front and back surfaces of the lens. These are visibly revealed by changes in the size and position of reflections from these surfaces known, respectively, as the 3rd and 4th Purkinje images. Although the 3rd image is difficult to produce in a measurable form (and to find even then), these reflections have been observed by many investigators and measured by a few including Wulfeck (1952) who took motion pictures of the reflections from two infrared point sources.

The recent explosion in accommodation research, however, has been made possible (and greatly stimulated) by the development of the Crane-Cornsweet three-dimensional eye tracker, an infrared oculometer/optometer combination available from SRI International (Cornsweet & Crane, 1970; Crane & Steele, 1978). Briefly, the device provides a continuous high-bandwidth output of changes in optical refraction required to keep an infrared image in focus on the retina as the eye accommodates. Its major limitation is that its measurements are relative to an approximation of zero diopter and therefore not absolute.

FUNNY THINGS THE EYE DOES

Since the original Crane-Cornsweet optometer was developed for NASA's Ames Research Center, the recordings of its outputs from many experiments have been full of surprises. The eye does some strange things; it is not only lazy and disobedient but also stubborn, emotional, and occasionally a practical joker. Several of these curious things can be seen in Figures 2, 3, and 4 which are based on representative stripchart recordings taken by Lynn Olzak and Donna Miller in an experiment conducted at Ames Research Center (Roscoe, et al., 1976; Roscoe, 1977).

These particular recordings were taken to determine the resting, or "open loop," accommodation level of each subject prior to his participation in the main experiment. Following a brief interval of preaccommodation to either a O D or 4 D stimulus target, a 1-mm diameter artificial pupil was positioned 8 cm in front of the subject's left eye while the subject continued to fixate the target. The artificial pupil serves to maintain a focused image as accommodation drifts " open loop" toward its resting level.

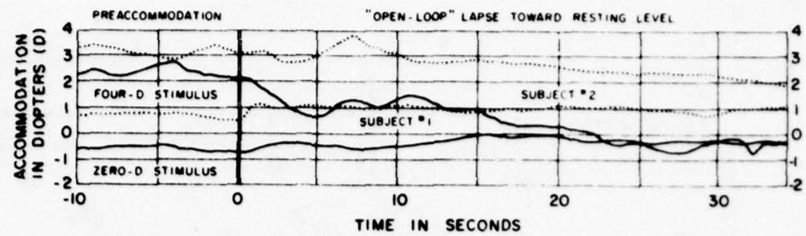


FIGURE 2. Preaccommodation to zero-diopter and four-diopter stimulus targets by two pilots, followed by their "open-loop" responses subsequent to the insertion of an artificial pupil in front of the left eye at time-zero.

Preaccommodation

In Figure 2, the recordings for two subjects have been smoothed to illustrate more clearly a number of typical findings: (1) Individuals accommodate differently to the same stimuli, particularly as their distance increases; to the O-D target, Subject 2 preaccommodated to a dioptric level that corresponds to a distance of 1-1½ meters, whereas Subject 1 preaccommodated to a distance beyond minus 1½ D (a response analagous to that of a zoom lens). (2) Subjects preaccommodated more steadily to more distant targets. (3) Despite the large displacement of preaccommodation levels for the two subjects, the differences in responses to O-D and 4-D stimuli by each subject were of a similar magnitude.

Open-Loop Responses

Upon insertion of the artificial pupil, Subject 1's eye took off in a hurry, wandered a bit, and then proceeded to its resting level just beyond optical infinity (< 0 D); Subjects 2's eye went immediately to 1 D from its preaccommodation to the O-D target and hugged the 1-D line thereafter; but, from its 4-D target, it initially ndered around as if confused, actually rising to almost 4 D at the 7-sec point, and then slowly lapsed toward a resting level of 2 D, showing a hysteresis of 1 D relative to its resting level from preaccommodation to the O-D stimulus.

These responses are typical of two types of subjects between and beyond which there is continuous variation within the normal population. Individuals (e.g. Subject 1) with distant resting accommodation tend to underaccommodate to either near or far targets and to lapse quickly to their resting level from either direction. Individuals (e.g. Subject 2) with near resting accommodation tend to overaccommodate to far targets and to lapse slowly and uncertainly from their near preaccommodation level toward a resting level that is much nearer than their resting level following preaccommodation to a far target. These are the individuals who contribute to the typically observed group hysteresis in resting accommodation levels.

Hyperopic individuals (like Subject 1) are readily trained to accommodate voluntarily, and when so trained, their relatively smooth spontaneous fluctuations in accommodation, to be seen in Figures 3 and 4, become even smoother, particularly at their resting distance. In contrast, myopic individuals (like Subject 2) are resistant to conditioning of accommodation control, and their "noisy" spontaneous flucutaions and frequent blinks persist. People like Subject 1 are referred to by Randle (personal communication) as "sympathetic" types: outgoing, flexible, attentive to their environment; those like Subject 2 as "parasympathetic" types: inward-looking, defensive, perservative. So much for Randle as a clinical psychologist.

And so much for typical findings; now for the surprises. Figures 3 and 4 follow the open-loop responses of the same two subjects for another minute and a half beyond the limit of Figure 2. These figures include the actual unsmoothed output of the optometer, as well as the smoothed curves, and show the spontaneous fluctuations in accommodation that range over about + 1½ diopter for Subject 2. These fluctuations are not unlike the spontaneous fluctuations in the line of sight that are essential to normal vision (Pritchard, 1961), and they suggest a mechanism that allows us to see clearly though not accurately accommodated by recourse to a scanning process.

But the big surprise is yet to come. In most tests of resting accommodation, responses have been recorded for only one minute. When Lynn Olzak asked the senior author how long to record the subjects' open-loop responses, he said, "Let's let it run for two minutes. Maybe the hystersis will wash out." The hysteresis did not wash out, but something else happened. At varying times during the second minute of open-loop response the eyes of several of the subjects entered a hunting mode of one type or another—as if they were looking for an out-of-focus image to back away from.

In the case of Subject 1, the hunt did not start until the fourth half-minute and searching strategy was oscillatory. For Subject 2, the strategy was different; after preaccommodation to a O-D target, this subject maintained a steady 1-D resting level for almost a minute, then shifted inward about 1½ D for about 40 sec, and then abruptly inward to the 2-D level, which he tended to maintain for the rest of the second minute. After preaccommodating to the 4-D target, this subject's eye drifted to a resting level of about 2 D within half a minute, wandered around between 2 D and almost 3 D for the next minute, then abruptly jumped to 3-1½ D and wandered near that level for the next half minute.

The responses just described are not unusual; in fact, they have been selected as typical of two different types, tending toward but not reaching the extremes. What they illustrate in common is that, while different eyes employ different strategis, eyes in general seek an out-of-focus image to back away from, thereby maintaining a sufficiently clear image for the task at hand. What else these recordings indicate is that the literature of visual perception of size and distance is replete with data that do not mean what the investigators thought they meant and a host of conclusions that can be dead wrong.

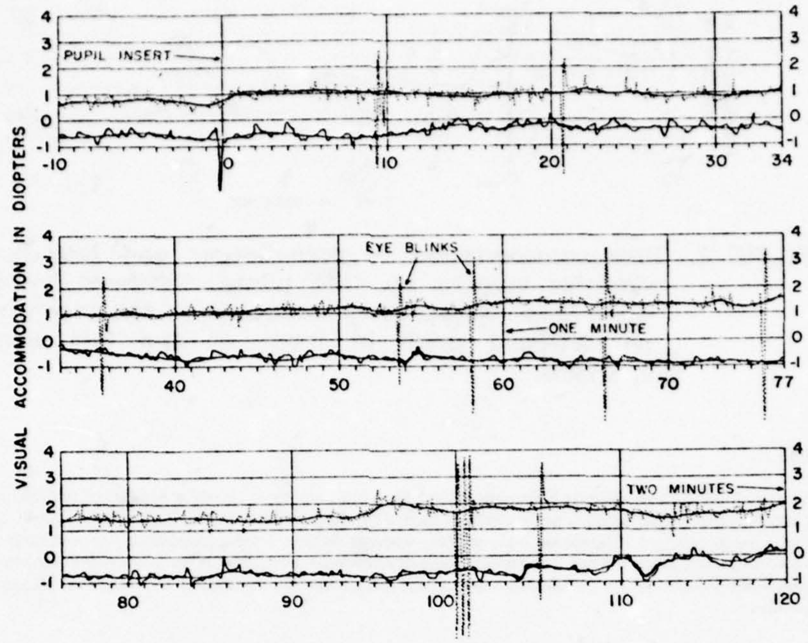


FIGURE 3. **Continuous records of "open-loop" accommodative responses of two subjects, with smoothed curves superposed, for two minutes following preaccommodation to a zero-diopter stimulus.**

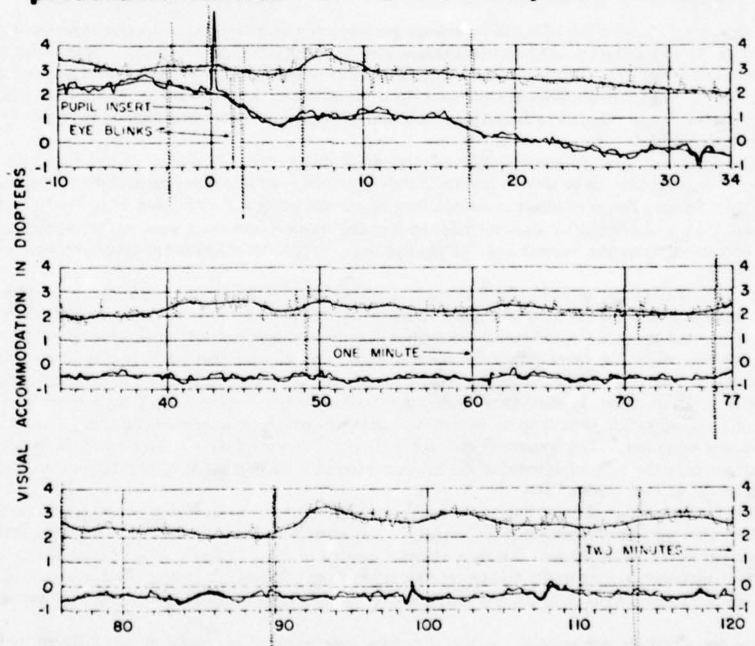


FIGURE 4. **Continuous records of "open-loop" accommodative responses of two subjects, with smoothed curves superposed, for two minutes following preaccommodation to a four-diopter stimulus.**

REFERENCES

Cornsweet, T.N. and Crane, H.D. Servo-controlled infrared optometer. *Journal of the Optical Society of America*. 1970, *60*, 548-554.

Crane, H.D. and Clark, M.R. Three-dimensional stimulus deflector. *Applied Optics*, 1978, *17*, 706-714.

Crane, H.D. and Steele, C.M. Accurate three-dimensional eyetracker. *Applied Optics*, 1978, *17*, 691-705.

Davson, H. *The physiology of the eye*. New York: Academic, 1972.

Duke-Elder, W.S. *Textbook of ophtahalmology, Vol 1*. London: Kimpton, 1940.

Helmholtz, H. von. *Physiological optics*. New York: Dover, 1867/1962.

Hennessy, R.T. and Leibowits, H.W. The effect of a peripheral stimulus on accommodation. *Perception and Psychophysics*, 1971, *10*, 129-132.

Holway, A.H. and Boring, E.G. Determinants of apparent visual size with distance variant. *American Journal of Psychology*, 1941, *54*, 21-37.

Leibowitz, H.W. and Hennessy, R.T. The laser optometer and some implications for behavioral research. *American Psychologist*, 1975, *30*, 349-352.

Leibowitz, H.W. and Owens, D.A. Anomalous myopias and the intermediate dark focus of accommodation. *Science*, 1975, *189*, 646-648.

Pritchard, R.M. Stabilized images on the retina. *Scientific American*, 1961, 204(6), 72-78.

Randle, R.J. Volitional control of visual accommodation. *Proceedings of the AGARD Conference on adaptation and acclimitisation in aerospace medicine*, Nuilly-sur-Seine, France: North Atlantic Treaty Organization, 1971.

Randle, R.J., Roscoe, S.N., and Petitt, J. Effects of accommodation and magnification on aim-point estimation in a simulated landing task. Washington, D.C.: National Aeronautics and Space Administration, NASA Technical Note (in press).

Roscoe, S.N. How big the moon, how fat the eye? Savoy, IL: University of Illinois at Urbana-Champaign, Aviation Research Laboratory, Technical Report ARL-77-2/AFOSR-77-2, 1977.

Roscoe, S.N., Olzak, L.A., and Randle, R.J. Ground-referenced visual orientation with imaging displays: Monocular versus binocular accommodation and judgments of relative size. *Proceedings of the AGARD Aerospace Medical Panel specialists meeting, Athens, Greece*. Nuilly-sur-Seine, France: North Atlantic Treaty Organization, 1976.

Whiteside, T.C.D. *The problems of vision in flight at high altitude* (AGARDograph No.13). London: Butterworths publication, 1957.

Wulfeck J.W. New techniques for an experimental analysis of accommodation time. Rochester, N.Y.: The University of Rochester, Doctoral dissertation, 1952.

Young, F.A. The projection of after-images and Emmert's Law. *Journal of General Psychology*, 1948, *39*, 161-166.

Young, F.A. Studies of the projected afterimages: III. The projection over large distances. *Journal of General Psychology*, 1952, *47*, 207-212.

HFE FOR THE PATRIOT MISSILE SYSTEM

N. Johnson, P. Williams, G. Kurtz, J. Gombash, and R. Camden
US Army Human Engineering Laboratory

INTRODUCTION: During the last 3 years an effort was undertaken by the US Army Human Engineering laboratory to develop a command control simulation facility for the purpose of evaluating a variety of display and control concepts and designs which are in the developmental stage and form a vital part of today's complex military systems. The PATRIOT display and control subsystem is one of those systems currently being studied.

The PATRIOT missile system (VG #1) is a new-generation air defense weapon system now in development. The unique features of the system are a multifunction phased array radar, the track-via-missile guidance system, and automated operation with human control.

This next picture (VG #2) shows the major equipment items comprising a PATRIOT firing platoon. The firing platoon consists of 5 major items: the radar set, the engagement control station, and the electric power plant constitute the fire control section, while five launching stations each including 4 missiles constitute the launching section. I should emphasize at this time that once the system is emplaced and operating, only the engagement control station is manned during air defense missions; all other items are remotely controlled and unattended.

This picture (VG #3) shows a view of the engagement control station which is where human control is exercised. Stored computer programs, as modified by operator selections and instructions inputted on either of the two operator consoles, control the entire system operation.

This photo (VG #4) shows the contractor's display and control consoles and this photo (VG #5) shows one of the two consoles which have been simulated by the HEL. Since they are universal consoles, either console can be used by an operator to perform all functions or the functions can be divided between the two consoles in various ways; e.g., engagement control on one console and battalion interface and status monitoring on the other console.

Briefly, the HEL simulation facility (VG #6) equipment consists of a Varian 620/F-100 minicomputer with 32K memory, a real-time clock and disc storage capability, plus an IDIIOM graphic display system with up to 4 cathode ray tubes, function keyboards, light pens, and tracking joysticks. Peripheral equipment includes a printer, teletype, and card reader.

The real-time simulation program uses the preprocessed scenario data to present a realistic display of aerial targets to the console operator. The target data interact with the operator's actions to simulate the total air-defense system. Operator actions and target events are recorded in real-time for later analysis.

Computer programs have been developed to aid in evaluating operator performance by isolating performance parameters and summarizing operator actions. There are also provisions for scenario replay with operator actions, inidivdual target history, chronological event listing, target kill-assessment summary with intercept locations, asset boundary-penetration summary, and switch-action summary.

Targets are identified and classified as they enter the radar-tracking sector, and are displayed as coded symbols. Each target is then subjected to various tests which determine its engagement eligibility and threat eligibility. If a target meets these criteria, further tests determine which asset it threatens. Then a launch-decision process ranks the target for engagement; this process if based on location of the predicted intercept, time urgency to protect an asset, and the initialized operator-time requirements.

The operator sees a basic picture like the one shown in this viewgraph (VG #7). This is a plan position indicator display which is on at all times. It shows the radar-boundary limits for searching and tracking targets, and range rings used as guide for target distance. Through the operator's switch actions, additional map data are available for display. These data include radar-masking terrain, air corridors, prohibited areas, restricted areas, defended areas, and the forward edge of the battle area (FEBA).

Dynamic elements shown in this next picture (VG #7A) appear and are updated under program control. These elements include target symbol modifiers, track numbers, velocity vectors, (VG #7B) defensive-missile symbols, launch-now-intercept lines, target-path history, and predicted intercept points. The operator has the option of displaying or removing these elements by switches, provided on the console panel.

Tabular displays appear in an area directly below the PPI graphic display. They consist of an alert-message line which is not shown in this picture and three mutually-exclusive tabular data displays which are a missile inventory, an engagement data display which is divided into to-be-engaged and engaged tracks and a track-amplifying data display. The next picture (VG #8) shows a complete rather active display as an air-defense operator would typically see it. An alert message is shown in this picture.

The console operator uses an isometric joystick to control a PPI-display cursor as a pointer to select targets for further evaluation and or engagement.

I will describe the operator performance testing performed on HEL's PATRIOT display and control simulator.

Since there was no test data available where PATRIOT operator performance could be compared, it was decided to obtain performance measures using the current display and control system design to establish a baseline of performance measures. Once this was done, software modifications were made to the design and tested to determine whether they improved or degraded operator performance over the baseline design.

The objectives of the operator performance testing were: (VG #9).

1. To evaluate the current PATRIOT display and control design for performing air-defense missions in a benign environment.
2. To develop display and control modifications for improving the current display design.
3. To determine the effectiveness of these modifications on air-defense operator performance.
4. To assess the adequacy of an HEL-developed tactical standing operating procedure for PATRIOT.
5. To examine operator performance under different target densities.

During our testing on the baseline system, we found that operators were having difficulty in performing the following tasks: (VG #10).

a. Detecting and engaging air-to-surface missiles within the short time period that these targets could be engaged and intercepted.
b. Remembering to move the tabular display cursor back to the line of the to-be-engaged tabular display after hooking a target on the 1st line.
c. Comparing TSOP hostile criteria with an unknown track's parameters for the purpose of declaring an unknown track hostile.
d. Remembering to select ripple method of fire on a multiple hostile track.
e. Detecting and rapidly responding to visual alert messages--even critical alert messages.

To reduce or eliminate these problems in hopes that operator performance would be significantly improved, the following software changes were made:

a. To improve performance on both detection of air-to-surface missiles (ASMS) and critical alert messages (problems a & e), ASMS which were not originally threat ordered were threat ordered in part: i.e., those ASMS which were targeted within a certain distance of the firing platoon were threat ordered in order to improve the self defense posture of the firing platoon. Thus, the ASMS which we threat ordered were now put on the to-be-engaged section of the tabular display along with other aircraft threats. When one of these targets met a certain criteria, it initiated a blinking priority engagement visual alert message to the operator along with an audible alert. With this change, all the operator had to do was to acknowledge that alert and issue an engage command.

In the baseline system there was no audible alarm provided for a critical threat ordered aircraft or ASM. At times the operator was not aware of the blinking alert message and had to depend strictly on his own surveillance capability to detect an ASM symbol on the PPI display. Once the ASM was detected, he had to use his joystick to manually hook the ASM track before he could issue an engage command. With the change, the ASM track was automatically hooked when the priority engagement alert message was acknowledged.

b. To eliminate the operator task of moving the tabular cursor back to the 1st line of the TBE tabular display, all that was needed was software change to prevent the cursor from moving down to the next line after a sequence hook action was taken by the operator. Notice that the next logical action an operator would take, after sequence hooking a track on the to-be-engaged display, would be an engage action. Once an engage action was taken, the track on the TBE display moved to the engaged part of the display and all other tracks on the TBE display were reordered with the track on the 1st line having the highest priority; however, the tabular cursor was now on the 2nd line and the operator had to move it back to the 1st line so that he could hook and engage the highest priority target. Remembering to do this task frustrated a number of subjects to a point where they completely refused to use the sequence hook function. I might add that once the change was made, the subjects found that this method of hooking was very effective, especially in times when there were many hostile tracks on the display and they preferred this hooking method over the manual/joystick hooking method.

The next picture (VG #11) shows the test results where our modification showed a significant improvement in operator/machine/mission performance.

PATRIOT DISPLAY AND CONTROL EVALUATION TEST RESULTS

DEPENDENT VARIABLES	BASELINE SYSTEM	MODIFIED SYSTEM
Unknown Track ID Time (sec)	50.6	33.4*
Priority Alert to Acknowledge Time (sec)	5.0	2.5*
Priority Alert to Engage Engage Time * (sec)	6.0	4.2*
Mean Time to Engage FP Directed ASMA* (sec)	15.0	10.0*

**Significant at .05 confidence level.*

**Significant at .01 confidence level.

A total of 16 military air defense personnel having varying amounts of operator experience on the HAWK air defense system served as subjects for the PATRIOT test.

The benefits which can be realized from this effort are:

1. It reduces the need and costs of conducting special human engineering field tests on the actual system equipment.
2. It provides visibility on human engineering problems existent in the current system design. Detecting and correcting these problems now is less costly than trying to correct them when the system is fielded.
3. It provides a technical Baptism to personnel in the test and training communities whereby they are better prepared for structuring their respective tests and defining their training requirements; HEL is also participating in developing a PATRIOT operator tactics trainer.

Finally, in regards to our future tests planned for the PATRIOT simulator. The topics to be studied are: (VG #12). *Future Tests Planned*

1. Display Study on Operator Capability in an ECM Environment.
2. Task/Workload Divisions Using 2 Display and Control Consoles.
3. Evaluation of a New Engagement Data Tabular Display.

As an added attraction which shows an aspect of the cost/effectiveness of using a simulator instead of range, actual missiles and aircraft I'll show you a 5-minute contractor prepared film on PATRIOT.

TRANSFER OF LANDING SKILL AFTER TRAINING WITH SUPPLEMENTARY VISUAL CUES

Gavan Lintern and Stanley N. Roscoe
University of Illinois at Urbana-Champaign

An aircraft simulator with a closed-loop computer-generated visual display, was used to teach flight-naive subjects to land. A control training condition in which subjects learned to land with reference to a skeletal airport scene consisting of a horizon, runway, centerline, and aiming bar, was tested against training with constantly augmented feedback, and a flightpath tracking display. A simulator to simulator transfer-of-training design showed that adaptively trained subjects performed best in a transfer task that was identical to the control group's training condition. Several subjects attempted six landings in a light airplane after they had completed their experimental work in the simulator. They performed better than another group of subjects that had not had any landing practice in the simulator.

INTRODUCTION

A simple and relatively inexpensive computer-generated visual display was used to teach flight-naive subjects to land an aicraft simulator. One aim of this experiment was to test augmented feedback techniques for simulator landing instruction. Another aim was to examine the value of a visual landing display, in conjunction with a ground-based trainer, for teaching airplane landing skills.

A review of the augmented-feedback literature indicated that perceptual-motor skill instruction could be enhanced with supplementary cues that provided more precise information. However, the literature also indicated that transfer to nonaugmented-feedback conditions would be disrupted if the subject learned to depend on the supplementary cues at the expense of the intrinsic cues. Depending on supplementary cues appeared to disrupt transfer performance particularly with poor-intrinsic-feedback tasks.

For this experiment, an off-course, adaptive augmented-feedback manipulation that did not permit dependencies on the supplementary cues to develop, was compared with a constant augmented-feedback manipulation. In addition, both augmented-feedback training conditions were compared with a nonaugmented-feedback training condition. Relative effectiveness of the various training conditions was assessed by comparing performances in a transfer condition identical to the nonaugmented training condition.

Previous research with visual landing displays has generally failed to control for the possibility that any performance advantage shown by subjects trained with the visual landing display may have been entirely due to additional control experience in the simulator to which the visual display was attached. A flightpath tracking display, in which subjects practiced control movements identical to those required for landing the simulator but with our reference to any representation of visual cues found in a normal airport scene, was used as a fourth training condition. Transfer performances after training with this condition were compared with transfer performances after training with the nonaugmented visual landing display.

Transfer of landing skill to the airplane was assessed by comparing airplane landing performances of four subjects who had not received any landing instruction in the simulator with those of the 21 subjects who had completed their simulator landing trials.

METHOD

1. *Procedure*

Four groups of 12 randomly allocated, flight-naive subjects were instructed in airplane systems, aerodynamics and flight technique, and were then taught non-contact flight maneuvers in a modified Singer-Link general aviation trainer with a visual display (GAT/VD), over three sessions. In a fourth session, subjects received approach and landing instruction in the GAT/VD under one of four experimental training conditions.

The fourth session was the only one in which instruction differed for the four groups. The *Primary Control Group* practiced landing the GAT/VD by reference to a closed-loop computer-runway outline with a centerline and an aiming bar. A second group, designated the *Constant Augmented-Feedback Group*, practiced landing with a visual display that had, in addition to the previously mentioned airport graphics, command guidance cues that showed the subject the desired flight path during the approach and the proper height during the flare. An *Adaptive Augmented-Feedback Group* practiced with the same display except that the command guidance cues appeared only when the subject deviated from the specified performance criteria. A secondary control group, hereafter referred to as the *Flightpath Tracking Group*, practiced a compensatory tracking task by reference to a fixed cross and a moving square projected onto the screen in front of the GAT/VD. Students had to progress through a descent, leveloff, and stall to perform the tracking task successfully.

A transfer condition identical to the *Primary Control* training condition was used to test simulator landing performances of all subjects in the fifth and any subsequent simulator sessions.

One subject from each group attempted six landings in a Piper Cherokee Arrow between their third and fourth sessions. Another 21 subjects attempted six landings in the same airplane after they had completed their simulator landing trials.

Roll response of the GAT/VD was turned off throughout landing trials. Pre-testing had indicated that the GAT/VD is sufficiently unstable about the roll axis to make it considerable more difficult to control than a light airplane, so much more difficult in fact that, even after training periods of four hours, control by beginning students was too erratic to provide experience with visual relationships typical of normal landing approaches. With the roll motion turned off, or "frozen", the difficulty of landing the GAT/VD more closely approximated the difficulty of landing an airplane.

2. Apparatus

The flight characteristics of the GAT/VD were modified to approximate those of a Piper Cherokee Arrow. Its motion system was switched off throughout the experiment. The visual display system consisted of an Advent Videobean 1000 television projector mounted on the cockpit roof of the simulator and a 1.78x1.35 meter (70x53 inch) spherical section screen placed approximately 2.24 meters (88 inches)

in front of the subject's viewing position. Dynamic imagery, generated on a 21.5 mm (8.5 inch) square plasma-matrix display panel by a PDP-11/40 computer was transferred to the television projector via a Panasonic WV-350P television camera. Following Roscoe, Hasler and Dougherty (1966), the simulated airport scene was magnified by a factor of 1.25. The runway was 1372 meters (4500)feet) long by 61 meters (200 feet) wide. The approach aiming bar was a double line across the runway, 152 meters (500 feet) from the threshold.

Films and audiotapes covering basic aerodynamics and flight maneuvers were loaned by Jeppesen-Sanderson for the ground school instruction. A Norelco model LCH 2020 variable motion projector, loaned by Phillips Audio Video Systems, was used to project and play the films and tapes.

RESULTS

Performances were assessed by scoring the number of criterion trials during the pretransfer phase and in the first 16 transfer trials. A criterion trial was one in which the subject remained within the flight envelope used to switch the supplementary cues for the adaptive training condition, from 914 meters (3000 feet) before the runway aiming bar to 305 meters (1000 feet) past it. Both augmented feedback conditions produced better pretransfer performances than the *Primary Control* condition, with the *Adaptive Augmented-Feedback* condition producing the best overall pretransfer performances. These trends were not statistically reliable.

In transfer, the *Adaptive Augmented-Feedback* Group outperformed both the *Constant Augmented-Feedback Group* and the *Primary Control Group* (p.01). The *Constant Augmented-Feedback Group* and the *Primary Control Group* demonstrated approximately equivalent performances. The transfer performances of the *Flightpath Tracking Group* was poorer than that of the *Primary Control Group*, although this trend was not statistically reliable.

Number of unassisted landings and subjective performance ratings by the safety pilots were used as dependent variables in the analysis of airplane landing trials. The group that landed the airplane after their simulator landing instruction performed better on both of these indices (p .05).

DISCUSSION

The superior transfer performance of the *Adaptive Augmented-Feedback Group* indicates that this type of manipulation can enhance landing training. Much of the research effort within engineering psychology into perceptual-motor skill instruction has been focused on adaptive training, a procedure in which the training schedule is individualized by varying task difficulty through a graded series of steps at a rate that is related to the student's speed of learning (Kelley, 1969). A demanding task is initially simplified in some manner, and control performance is monitored so that the task difficulty or complexity can be gradually increased as the student develops proficiency. Adaptive training is based on the intuitively appealing assumption that a demanding task can be learned more efficiently if it is presented throughout training at a level of difficulty that is optimally matched to each individual's current ability.

Although adaptive training has generally found enthusiastic support within engineering psychology and in some applied settings, there are few data that attest to its advantages for teaching vehicular control skills. An extensive research literature has generally failed to demonstrate that adaptive manipulations of system variables, such as order, gain, or lag and input variables such as amplitude or frequency or a forcing function, can speed acquisition of control skills. Data trends have generally shown no advantage for the adaptive manipulations, and several reports of positive results do not, under closer examination, support the application of adaptive training (Lintern and Gopher, 1977). Nor do they clearly demonstrate that adaptive training cannot be effective.

The off-course augmented feedback manipulation used in this experiment conformed to the basic requirements of an adaptive manipulation. It simplified the task, as shown by the superior performance of the *Augmented-Feedback Group* during training, and was inherently adaptive in that, as the student improved, the augmenting cues that simplified the task appeared less frequently. The process was one of automatically withdrawing the supplementary cues at a rate that was determined by individual improvement. Lintern and Gopher (1977) have argued that augmented feedback is likely to provide the most effective adaptive training manipulations of those yet tested. The results of this experiment support that argument.

The hypothesis that constant augmented feedback would not assist transfer performance, because subjects would become dependent on the supplementary cues, was not closely supported even though the *Constant Augmented-Feedback Group* did no better than the *Primary Control Group*. This result could be attributed to the inability of the cues, when appearing throughout the landing trial, to appreciably aid pretransfer performance.

The *Primary Control Group* demonstrated better transfer performance than the *Flightpath Tracking Group*. This trend conformed to the belief that merely practicing the correct control responses is not as effective as practicing the correct responses with feedback from a closed-loop visual display. However, subjects in the *Flightpath Tracking Group* appeared to make more control reversals than did subjects in the *Primary Control Group*. It is possible that incompatible control-display relationships disrupted learning in the flightpath tracking condition. Thus the comparison of the flightpath tracking and primary control transfer performances failed to resolve the importance of a closed-loop visual landing display for simulator landing training.

The primary purpose of the airplane landing trials was to provide a validity check on the GAT/VD. With the roll response frozen and severly limited detail in the visual display, there was some doubt whether the landing practice in the GAT/VD would teach any skill that could be used to land an airplane. Within the limitations of the test that was undertaken, the simulator landing had a strong and statistically reliable influence on the airplane landing work. It helped students both in indices of unassisted landings and on subjective instructor ratings of performance. It is encouraging to see that this fairly rudimentary device has demonstrated worthwhile transfer to the airplane so early in its development.

ACKNOWLEDGMENT

This research was conducted at the Aviation Research Laboratory, Institute of Aviation, with contractual support from the Life Sciences Program, Air Force Office of Scientific Research, Contract Number F44620-76-C-0009. Dr. Alfred R. Fregly was the scientific monitor of the contract. Films and tapes for ground school instruction were loaned by Jeppesen-Sanderson, Denver, Colorado, and projection equipment was loaned by Phillips Audio-Video Systems Corporation, Mahwah, New Jersey.

REFERENCES

Kelley, C.R. What is adaptive training? *Human Factors,* 1969, *11*, 547-556.

Lintern, G. and Gopher, D. Adaptive training of perceptual motor skills: Issues, results and future directions. Savoy, Ill.: University of Illinois at Urbana-Champaign, Aviation Research Laboratory, Technical Report ARL-77-5 AFOSR-77-5, 1977.

THE DEVELOPMENT AND ANALYSIS OF THE F-15 FLIGHT SIMULATOR COMPUTER SCORING ALGORITHM

David W. Evans, Bruce A. Smith, Michael J. McDonald
United States Air Force Academy

INTRODUCTION

Pilot performance is a subject of great interest in the United States Air Force. The whole concept of pilot performance assumes that there are differences in pilots and there are measurable characteristics that distinguish these differences Fessler, (1977). There are two major schools of thought concerning how to measure these differences. One school, the traditional method, emphasizes the check pilot evaluator and completely subjective measures. The second school, objective, emphasizes the factual, non-emotional machine type scoring which is directly applicable to simulator flying Fessler, (1977).

Research done in the subjective methods of scoring have shown that a number of difficulties exist. Although the human analytical approach has the advantages of human involvement and fully exploits human understanding and observation powers, it is time consuming Knoop, (1973). Still Koonce, (1975) and Obermayer and Vrurls, (1974) noted the importance of providing the human observer with objective standards. Johnson and Boots, (1943) pointed out the importance of objective measures for pilot prediction. This was further emphasized by Mahler and Bennett, (1950) and Ericksen, (1951) who also supported the need for objective measurement systems. They also indicated the importance of setting a baseline for comparative measurement. Regardless of how carefully designed and controlled a subjective measurement is, it would still experience some difficulty in lack of reliability and in inconsistent discrimination Fessler, (1977).

With this in mind the development of objective systems seems to be appropriate in order to obtain the most advantageous method of scoring. This leads to design philosophies that stress reliance on behavioral task analysis data Cream and Lambertson, (1975). There are, however, problems existing in these areas. One of the most critical problems is that the maneuvers that seem to be the best discriminators of performance are also extremely difficult to measure. Thus, the machines in the early seventies were somewhat below the technological level needed for these maneuvers Long and Varney, (1976). This problem has been alleviated by the advent of the modern, digital-computer flight simulators such as the Advanced Simulator for Undergraduate Pilot Training (ASUPT), the Simulator for Air to Air Combat (SAAC), and the F15 Simulator. Because of special capabilities of these new general simulators the concept of objective measurement is becoming a reality. It is necessary, however, to insure that the measurement criterion, baseline measures, and resulting performance scores are validated before applying them to operational situations.

Since the FS was designed to be an aid to the Instructor Pilot (IP)-student relationship, it is necessary for computer evaluations to correlate with IP evaluations of the student's performance. In this light, measures must not only assess the degree to which objectives are met and the most salient characteristics of performance, but do so in such a manner that they agree with IP interpretation of the maneuver. This study was to fulfill that requirement, and to recommend adjustments in parameter scoring and parameter tolerances in order to improve correlations between IP and computer scores.

METHOD

In the early months of 1977, the Tactical Air Command (TAC) began an operational test and evaluation (OT&E) of the F-15 Flight Simulator (FS), built by Goodyear Aerospace, and located at Luke AFB, Arizona. AFHRL FTO at Luke AFB, was asked to assist in the conduct of the operational evaluation of the F-15 FS in several areas. This report deals with Objective #3 of the OT&E—Evaluate the student scorings algorithms, including parameters.

Subjects

Subjects in this study were all Air Force pilots engaged in the F-15 training program at Luke AFB, AZ. Pilots were from Nellis- and Luke-based F-15 squadrons, the 555th, 461st, and 433rd; each unit providing the instructor pilots (IPs) for their own subjects.

Subjects ranged in rank from First Lieutenant to Lt Colonel, their experience in aircraft ranging accordingly commensurate to rank.

Instructor Pilots were more uniform in rank, clustering around junior to senior captain. Their position testified to their excellence in flying. Individual methods of instruction and grading ranged along the continuum from the highly personal "good guy" to the professional "hard nose," although they were consistent within each IP. IPs participated in the program through altruism, gaining nothing directly from their work in the analysis.

Apparatus

The F-15 simulator at Luke AFB, AZ was the experimental device of study and analysis. Performance of the subjects in the simulator was graded by the IPs and the computer, providing a basis for comparison between the two. The simulator has a 6 degree of freedom motion base. However, it does not have a visual field other than the Heads Up Display. It was designed by the Goodyear Aerospace Division and is currently maintained by Goodyear and Air Force personnel.

The simulator has been programmed to evaluate the subject's performance in certain areas of flight. One departure from and 3 approaches to Luke and its auxillary field were selected for evaluation: Flatiron Departure, Luke HI-ILS 3R, ILS RW 11 JAY HI-TACAN. Although the computer recognizes 15 parameters, only 3 may be used on any one scoring leg. The departure and approaches have been broken into legs. The scoring algorithm grades deviation from and time out of the accepted standards. Points are then subtracted from 100, with -99 being the lowest possible grade.

Procedure

Since the analysis was to be done on the effectiveness of the existing system, initial parameters and standards of acceptance were presumed to be correct. A program of the algorithm was obtained and compared with flight maps to determine the description of the leg. A list

of the three most common parameters used throughout the algorithm was compiled. A scale was devised with 0 to 4 along a continuum, the middle region expanded to allow for greater differentiation among scores. The description, scale and parameters were then combined on AFHRL Form 135.

Each class of subjects had twelve periods of instruction, each period lasting one and one-half hours. The instructor pilot was briefed at the beginning of each period. He was explained the purpose of the study and what would be required of him. The experimenter reviewed the form after the period was over.

The first questionnaire section on solution of parameters was often completed hastily due to syllabus requirements. Therefore, a second questionnaire was prepared and distributed to the squadrons through their operations desk approximately two weeks later.

An average of the performance evaluations was calculated for each approach or departure. These were ranked and matched with their respective computer score. A Pearson Product Moment Correlation and a Spearman Rank Coefficient were then performed on the data. The results of the parameter questioning on both the first and second surveys were tallied. The nature of the expanded grading scale approximated interval data, therefore allowing for analysis with the Pearson's Correlation as well as the Spearman's.

RESULTS

The results of the evaluation are presented and discussed according to two topical areas: (1) the importance of parameters scored and (2) correlation of IP to algorithm scores, without modification.

Parameter Importance

Differences between parameters selected by the IPs and those currently in use by the scoring algorithms are negligible for the Hi ILS 3R, and JAY HI-TACAN approaches, and the Flatiron Departure. However, for ILS RWY 11 approach, the IPs suggested changes in 28 percent of the possible parameters to be scored.

Correlation of IP and Computer Evaluations

A significant positive correlation existed for only one of the mission profiles, Flatiron Departure 03. The approach profiles yielded nonexistant to negative correlations between IP and computer scores.

TABLE 1
CORRELATIONS BETWEEN COMPUTER AND IP SCORES

HI-ILS 3 R APP		FLATIRON 03 DEP	JAY HI TAC-APP	ILS 11 APP
Pearson	r = -.806	r = .752	r = -.102	r = -.533
Spearman	r = -.908	r = .639	r = .01	r = -.521
	N = 11	N = 26	N = 14	N = 13

Discussion

An apparent discrepancy exists in the results. While the computer and IPs use similar parameters in evaluating a student's performance, a significant positive correlation exists only for one of the mission profiles. This discrepancy should be corrected by improving simulator initializations for the approaches, and further research should be conducted before tolerance modification.

The major problem of the study, however, was intrinsic to the computer. Departures were all initialized from the same location, the runway, and therefore IPs and the computer started their evaluations at identical points in space and time. However, not all students entered the approach patterns at identical locations. IPs would often begin grading the approach before or after the computer's algorithms recognized the maneuvers as such. While the IP might find fault with the student's early maneuvers in the approach pattern, the computer would only grade the latter portion or vice-versa. Thus, an IPs evaluation would be lower/higher than the computer's, which did not begin its grading at the same point that the instructor did.

Recommendation

To facilitate future studies on the parameters, IPs should receive a pre-experimental briefing, clearly outlining instructions and objectives of the study. Close coordination should be done with the squadrons, and a scheduled time table drawn up that would allow maximum use of the FS by the squadrons and the research personnel.

Due to time limitations, two areas of possible inquiry were eliminated from this study, which might be useful in later research.

Prior to evaluating the computer program, a questionnaire should be prepared and sent to the squadron areas on which parameters should be scored. This would eliminate section 2 on the grade sheet, and remove busy work for the IP during the data collection phase of research.

In addition, future studies might consider revision of the syllabus in order to receive greater experimental opportunities. Several lessons were impossible to obtain data from due to heavy syllabus requirements. These lessons might be modified or ignored for data collection purposes.

Implications

As currently programmed, the FS is qualified for use in the departure mission profile. This pilot study should be followed by additional research to determine if discrepancies between IP and computer scores were due to: (1) parameter selection, or (2) difficulties encountered with the research.

In either case, due to unavailability of standardization-evaluation check pilots the FS *should not be used as a check ride instrument* on computer scoring alone, until extremely high positive correlations are obtained. The FS scoring program should aid students working together on Buddy Rides, or could be used in proficiency checks. The FS is a valuable tool, that with improvement, will aid instructors and students both.

REFERENCES

McDonald, Michael J., Evans, David W., Smith, Bruce A., Baer, Lester H. The Development and Analysis of the F-15 Flight Simultor Computer Scoring Algorithm, AFHRL, in press, Summer 1978.

NOTE: For a complete list of references see above AFHRL Technical Report.

SIGNAL ADJUSTMENTS FOR FORMATION FLIGHT IN THE ADVANCED SIMULATOR FOR PILOT TRAINING

Ricard, G. L., Cyrus, M. L., Cox, D. C.,
Templeton, T. K., and Thompson, L. C.
Naval Training Equipment Center, Orlando, Florida and
Air Force Human Resources Laboratory
Flying Training Division,
Williams Air Force Base, Arizona

During the past 20 years, digital controllers have replaced analog ones as the signal processors for flight trainers, mainly because of the accuracy and flexibility. With computer-driven visual displays just becoming "operational" for flight training, the switch to digital controllers is occurring now for flight simulator visual systems, mainly to take advantage of digital processors' second characteristic. Computer image generation (CIG) visual display systems cannot yet include all of the detail one might wish (because of their limited edge-displaying capacity), but they are the most flexible visual system yet developed for flight training. To accommodate different training tasks, for instance, the external visual world can be changed merely by placing new data into the display computer's memory. During the real-time simulation, this processor uses the x-, y-, and z-axis positions and the pitch, roll, and yaw angles of the simulated aircraft to adjust the information displayed for the position of the aircraft relative to the visual environment.

In such a visual system, a trade exists between the amount of detail in the visual display and the time required to perform the calculations necessary to adjust the visual image for movement of the simulated aircraft. The usual result is that the number of edges that can be presented is fixed, thus setting a processing duration. This time, added to that needed to process the flight dynamics, introduces a transport delay between a pilot's control inputs and the displaying of the aircraft's responses that presently exceeds 100 milliseconds. Reviews of the literature of manual control (Muckler and Obermayer, 1964; Ricard and Puig, 1977; and Poulton, 1974) indicate that tracking accuracy is reduced when a display delay is inserted, but the real problem associated with delays in CIG visual systems is "flutter" caused by adjustments made to the values of the simulated aircraft's pitch and roll angles for the time used by the CIG processor to update its image. Usually this adjustment is of the form $X_{n+1} = X_n = f(\dot{X}_n, \dot{X}_{n-1})$ where X is an aircraft parameter, $\dot{X}$ its derivative, the subscript refers to a value at a point in time, and f is a continuous function chosen to maximize lead information about the parameter X. Specifically, for the Advanced Simulator for Pilot Training (ASPT), values are adjusted according to the formula $X_{n+1} = X_n + H/2(3\dot{X}_{n+1} - \dot{X}_n)$. This amplifies the higher frequency components of the parameter X, which in turn are sent to the CIG system and displayed to the pilot as annoying flutter. More accurate prediction schemes can reduce this amplification, but some flutter will result when pilots use high-frequency control inputs to "fly" a simulator. We hypothesized that the information useful for aircraft control (and hence for meaningful simulation) is concentrated primarily in the mid- to low-frequency band and that above one to two Hertz, the relative amplitude of noise to *useful* signal increases. By following a prediction scheme with a low-pass filter, it should be possible to determine a trade, if it exists, between insufficient noise rejection (flutter) and information suppression (filtering too strongly). This we have attempted to do in a research simulator.

Usually, it has been flying tasks that require accurate control of a simulated aircraft that have proved difficult in a CIG system-equipped device, so to study the effect of filtering some of the signals sent to a CIG display, we used as an experimental task flying formation in the ASPT at Williams Air Force Base. The ASPT CIG system is a wide field-of-view display produced by a mosaic of seven 36-inch CRTs that surrounds each cockpit. It is capable of providing a visual scene covering 300° horizontal, and 150° vertical which duplicates that available in real aircraft. Larson and Terry (1975) have described the processing sequence of signals used by that display system, and Recard, Norman, and Collyer (1976) have presented data that indicate the training value of filtering some of these. In order to perform this suudy, low-pass filters had to be integrated into the ASPT visual-prediction routine, and the development of a way to do this that requires little processing time has been documented by Cyrus (1977).

METHODS

Using eight instructor pilots for the T-37 and T-38 aircraft as subjects, it was possible to do two experiments to determine a break frequency for the display-signal filters. The first was a 3 x 6 factorial using three conditions of aircraft motion cueing and six filter settings, and the second a simple comparison of the six conditions of filtering. The effect of adding nonvisual indicators of aircraft motion was studied by adding to a fly-with-just-the-visual-display condition, the cues provided by the motion base or G-seat of the ASPT. Constants were set in the filtering routine to produce -3 dB points of 1/4, 1/2, 3/4, 1, and 2 Hz and an unfiltered (normal display) condition for filters through which were passed the pitch and roll angles of the ASPT before they were sent to the CIG display system. Both experiments represented repeated measures designs and were analyzed by analyses of variance.

Our experimental task was to maintain the fingertip position off the right wing of a lead aircraft. For the first experiment, this aircraft flew an S-3 with no turbulence and during the second, it flew straight-and-level with moderate turbulence added. Flying performance was measured by recording differences of x-, y-, and z-axis position of the two aircraft and the differences of pitch, roll, and yaw angle. We also recorded pilot control inputs; these were the positions of the control stick, rudder pedals, and throttles. All of these were sampled at 15 Hz, and from these records, we calculated each measure's average value, variance, and bandwidth for every trial. At the end of each trial of both experiments, the pilots were asked to assign a rating to the "noisiness" of the CIG display and to the controllability of the simulator.

RESULTS

Experiment 1. As one might expect, we found significant trials and subjects effects as the instructor pilots were not particulary familiar with the ASPT, but we did see some effects specific to the filtering. The typical pattern of these is presented in Figure 1 where the variance of the difference of pitch angle of the two aircraft is presented across settings of the display filters. As the filter's break frequency was

made higher and higher (passing more and more low-frequency information), a reduction of the variance is seen until a setting of3 4 to 1 Hz is reached. Then, for this measure, a slight increase occurred when the filtering was removed (the ∞ break frequency condition). Similar patterns were recorded for the variance of the lateral and longitudinal movements of the control stick and for the average value of the throttle positions. Generally speaking, the measure most consistently affected by filtering was the variance of a parameter—a finding similar to results obtained by Cooper, Harris, and Sharkey (1975).

Figure 2 presents the ratings our pilots made of the controllability of the simulation of formation flight (those made of the noise of the CIG display system showed no effects), and here we found differences due to conditions of motion cueing and filter setting. The device appeared less controllable with its motion base activiated, and the best filter setting for flight control was judged to be 3 4 Hz, at least for the no-motion-base conditions. Because a fair amount of within groups variance was encountered, it was decided to make a replication a simpler version of the experiment.

Experiment 2. Instructor pilots familiar with the ASPT were used this time and a task was designed to emphasize the CIG-system noise. When this was done, the pattern of results seen in the first experiment was made clearer. Differences of pitch and roll angle were related to changes of the variance and bandwidth of the movements our pilots made with the control stick. The variance of the lateral and longitudinal movements of the control stick are depicted in Figure 3. Here is the now familiar decrease of variance until a setting of 3 4 Hz is reached with, this time, no evidence of an increase when the unfiltered-display condition was tested. Similar trends were found in the mean and variance of the positions of the rudder pedals and the variance and bandwidth of the throttle settings. Other significant effects were found, but we feel that they are probably either specific to this simulation of formation flight or are spurious.

Both sets of ratings showed significant trends across conditions of filtering, with the ratings of flight control paralleling the flying performance and measures of control input (as the pilots generated less error, they judged the device more controllable), and with those of the desirability of the CIG visual display decreasing as the signals were less and less affected by the filtering. These data are depicted in Figure 4. If Figures 3 and 4 are compared, it can be seen that a range of filter settings exists where pilots' opinions of the visual display are still tolerably high and flight control is little effected. This range we feel is 3/4 to 1 Hz.

DISCUSSION

The premise of these studies was the notion that only low-frequency information was necessary for flight control while it was the high-frequency components of the signals sent to the CIG display that were annoying. Clearly these can be separated. Flight control was not degraded (for the most part) by filter settings of 1/2 Hz or above, while at the same time, pilots' opinions of the display became lower and lower. The dramatic improvements of these opinions occurred at the upper end of our range of filtering conditions, so that a range (3/4 to 1 H can be defined where the display is acceptable and flight control is not affected. We feel that these data, collected in a device with a transport delay fixed by its architecture, confirm the results of laboratory studies—that changing the signals sent to a CIG visual display system can improve pilots' flying performance and their acceptance of a device so equipped.

REFERENCES

Cooper, F.R., Harris, W.T., and Sharkey, V.J. The effect of delay in the presentation of visual feedback on pilot performance. Technical Report NAVTRAEQUIPCEN IH-250, Naval Training Equipment Center, December, 1975.

Cyrus, M.L. A general method for compensating transport lags in computer image generation visual displays for flight simulation. Technical Report AFHRL-TR-77-6, Air Force Human Resources Laboratory/Flying Training Division, March, 1977.

Larson, D.F. and Terry, C. ASUPT visual integration technical report. Technical Report ASUPT -82, Singer-Simulation Products Division, March, 1975.

Muckler, F.A. and Obermayer, R.W. Control system lags and man-machine performance. Contractor Report NASA CR-83, National Aeronautics and Space Administration, July, 1964.

Poulton, E.C. *Tracking Skill and Manual Control.* New York: Academic Press, 1974.

Ricard, G.L., Norman, D.A., and Collyer, S.C. Compensating for flight simulator CGI system delays. In Proceedings of the 9th NTEC Industry Conference, NAVTRAEQUIPCEN IH-276, November, 1976.

Ricard, G.L. and Puig, J. A. Delay of visual feedback in aircraft simulators. Technical Note NAVTRAEQUIPCEN TN-56. Naval Training Equipment Center, March, 1977.

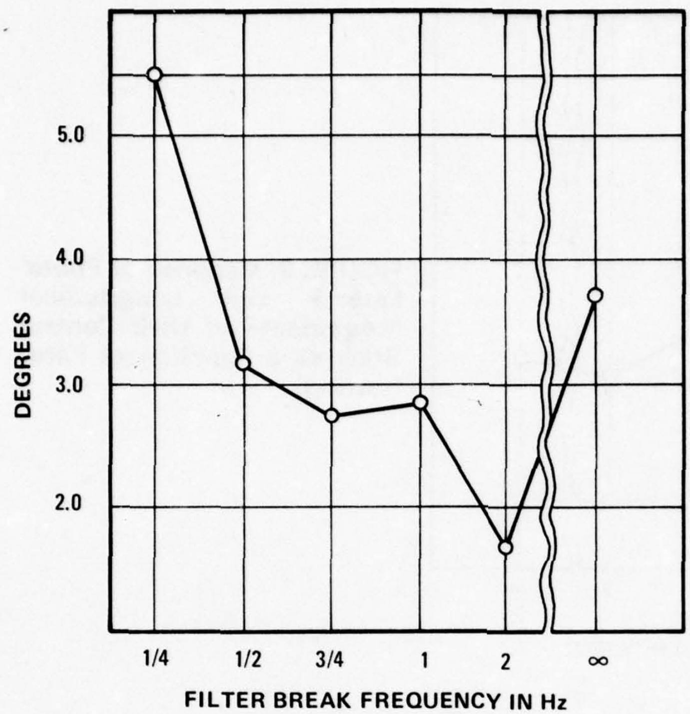


FIGURE 1. Variance of the Difference of Pitch Angle of the Lead and Lag Aircraft as a Function of Filter Setting. The ∞ symbol is used to indicate the no-filtering condition even though the Nyquist frequency of the ASPT CIG system is 15 Hertz.

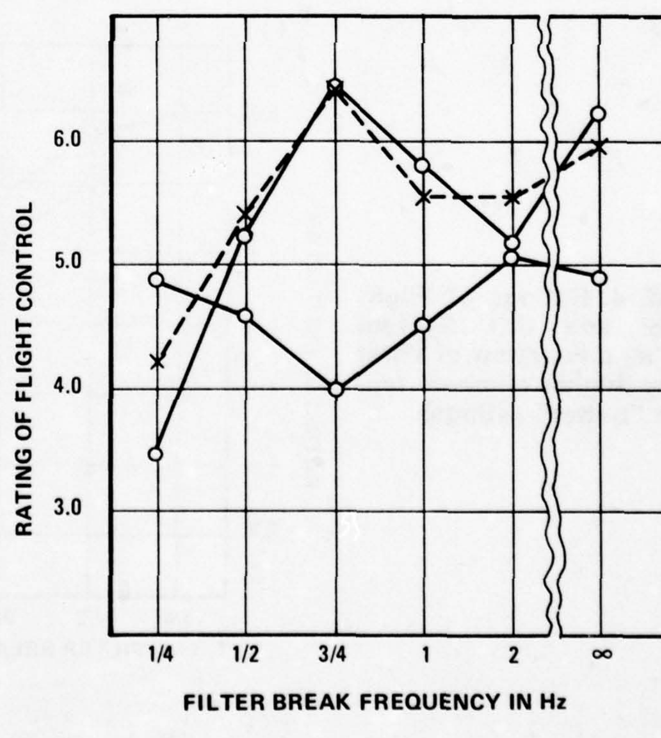


FIGURE 2. Ratings of Flight Control as a Function of Filter Setting and Conditions of Motion Cueing (high numbers represent "better" ratings and the cueing conditions are V=visual display only, V+M=visual display plus motion base, and V+G= visual display plus g-seat).

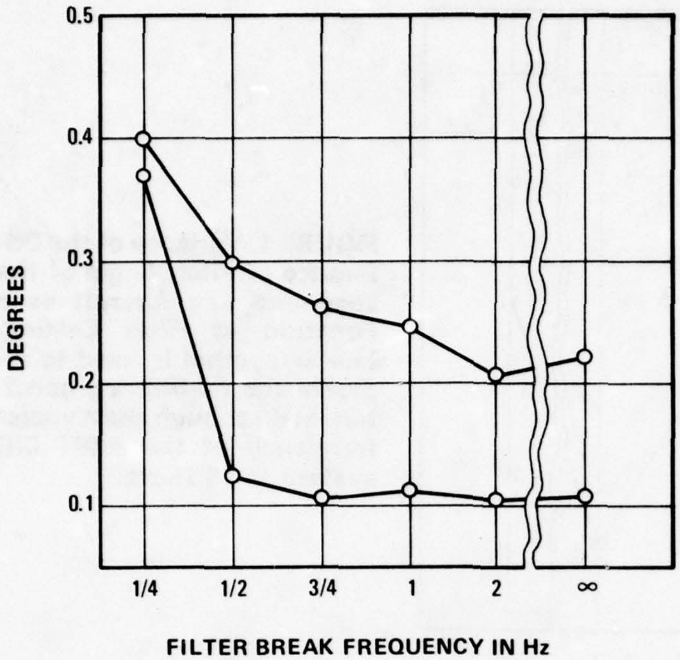


FIGURE 3. Variance of Pilots' Lateral and Longitudinal Movements of their Control Stick as a Function of Filter Setting.

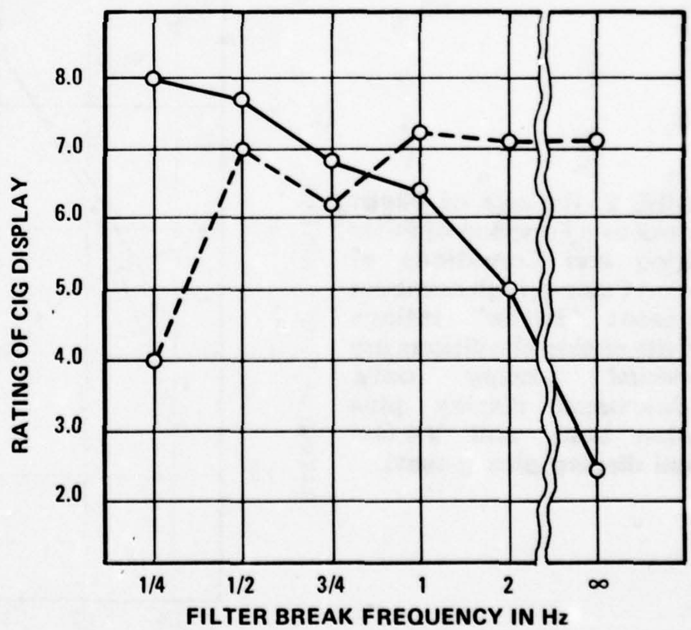


FIGURE 4. Ratings of Flight Control and CIG System Noise as a Function of Filter Setting (higher numbers represent "better" ratings).

PANEL DISCUSSION:
A Multi-Disciplinary View of Leadership in the Institutional-Occupational Framework

This panel discussed problems of contemporary leadership in the military in light of issues raised by Dr. Charles Moskos' model of Institutionalism versus Occupationalism including some essential differences in management philosophies and technique in an institutional environment versus an occupational environment.

Co-Chair: Lt Col Jock C. H. Schwank and Capt B. Don Sullivan, Department of Behavioral Sciences and Leadership, USAF Academy, Colorado.

Discussant: Colonel Thomas Wilkinson, USAF
Headquarters USAF
Washington, D.C.

Members: Dr. Charles Moskos
Northwestern University
Evanston, IL

Colonel Henry M. Kelly, USAF
Leadership and Mgmt Dev Center
Maxwell AFB, AL

Colonel John W. Williams, USAF
Dept of Beh Sci and Leadership
USAF Academy, CO

Colonel Malham Wakin, USAF
Dept of Philosophy & Fine Arts
USAF Academy, CO

Commander Dana French, USN
Bureau of Naval Personnel
Washington, D.C.

Lt Col David MacIssac, USAF
Department of History
USAF Academy, CO

Lt Col Howard Prince, USA
Dept of Beh Sci & Leadership
U.S. Military Academy, N.Y.

FROM INSTITUTION TO OCCUPATION: IMPLICATIONS AND NEW DIRECTIONS FOR THE U. S. ARMY

Lt Col Howard Prince, USA
U. S. Military Academy

By now Dr. Moskos' thesis concerning the movement of the all-volunteer military toward a more occupational model is well known in both military and academic circles and I will not elaborate further on the details of his arguments and analysis. Moreover, much of the evidence and support for Dr. Moskos' work on the shift of the military toward legitimation in terms of marketplace considerations was derived from careful study of the Army and, in particular, the first-term enlisted man. Such analyses have been penetrating and persuasive with respect to their reality, their present effects, and their possible organizational consequences such as military unions and increasing military plurality through greater reliance on civilian technicians performing military or quasi-military functions.

In what other ways might Dr. Moskos' hypothesis be explored? What are the implications of the ascendancy of an occupational model for Army leadership? What evidence might there be of trends toward an occupational model among other groups in the all-volunteer Army such as the NCOs and officers? What other factors might contribute to the apparent trends besides those already analyzed by Dr. Moskos? Finally, what are the implications for a professional military ethic of increasing drift toward the occupational model?

In looking at the first question I have raised, there are several observations which appear relevant. In terms of current Army leadership practices, there has been and continues to be increasing emphasis upon greater usage of civilian management theory and practices. The present Chief of Staff of the Army, General Rogers, has committed the Army to the training and use of internal organizational consultants skilled in organizational development. These school-trained officer consultants are assigned as staff officers to unit commanders in order to enhance organizational effectiveness. Implicit in their approach to assisting commanders in greater concern for subordinate (employee?) needs and interests. Is this shift in leadership practice merely a modern reemphasis of looking out for the welfare of your men, a reflection of the willingness of senior Army leaders to incorporate promising ideas whatever their source, or an organizational necessity in light of the trend toward an occupational model? What is the applicability of OD under conditions of increased readiness or possibly even combat? Does the Army need one kind of leadership for peacetime and another for combat in light of present trends? Finally, does the resort to civilian management practices actually *reinforce* the tendency toward an occupational model through increased recognition of individual interests and emphasis on the officer as manager and technocrat rather than heroic fighter leader?

In light of the arguments made in support of the occupational model hypothesis based on factors relating primarily to the enlisted man, one wonders what a careful look at the Army officer corps might reveal. Recent data collected by an Army officer for his Ph.D dissertation suggest that the Army may be less of a calling than one might expect among officers who represent the younger nucleus of the Army officer corps. Among regualr Army officers assigned to the staff and faculty of the US Military Academy who are by selection, in the top 1/3 of their contemporaries, it was found that their Central Life Interest scores (Dubin) were best characterized as having a flexible or diffuse rather than a work or job-centered focus. The CLI score has been found to be closely related to personnel turnover. What is the significance of CLI scores reflecting other than a strong job-centered focus for the military officer? Further research is needed to clarify this issue. What might be some possible sources or organization factors which contribute to these findings? Here we may speculate about new career realities for officers in the post-Vietnam all-volunteer Army and their possible effects. The Army may be engaging in officer personnel policies which contribute to the findings just discussed and which ultimately may lead to a shift from the calling and professional orientation of the Army officer. Continued emphasis on up-or-out promotion policies, successively lower selection rates for both temporary and permenent selection to the grades of major through colonel, the emphasis on increased specialization through the shift to OPMS, the commissioning of USMA graduates this year in other than combat specialties, and greater awareness of the realities of promotion to the general officer level (36 colonels selected for BG from among 4300 considered by most recent board) may individually or collectively contribute to an erosion of the Army Officer's commitment to a calling while increasing individual considerations and concerns. What behavior can we point to as evidence of the possible influence of such forces? High resignation rates among recent USMA graduates at the end of their initial commitment may be one such signal. Many young officers with an eye to keeping future options open take early advantage of off-duty education programs to acquire MBA degrees even after they have already been funded to earn advanced degrees in preparation for reduced promotion opportunities and increased competition for job security through increased selectivity for tenure in the form of permanent promotions follow with keen interest economic developments, especially recommendations concerning military retirement which may be viewed as deferred compensation for the hardships of an Army career.

A potentially related issue which impacts significantly on the future of military officership as a profession is that of the efficacy of the selection and socialization processes for new officers, particularly those trained by and commissioned from the service academies. Recent patterns of high attrition and repeated honor scandals leading to far-reaching changes at West Point are cause for concern in that the expensive service academy educations are justified on the basis of providing a unique socialization experience beyond the college education available through other commissioning sources. If the service academies are unsuccessful in either selection or socializing military novitiates what is the future of the calling in light of other trends toward an occupational model?

The foregoing leads me to the final related area which is what are the implications for a professional ethic in light of increasing drift toward the occupational model? Major Allan Futernick, a military sociologist who is currently a student at the Army CGSC, has focused on this issue in a recent provocative analysis. Briefly, Futernick points out that the role of the military in society is public service. This role ideally engenders a high degree of public trust accompanied by public expectation of adherence to high standards of ethical behavior. He goes on to contend that the shift from a calling to an occupational model holds within it potential for the development of an ethical basis inconsistent with the trust and expectations of the public. This follows from Futernick's analysis of the types of ethical codes the public is most likely to expect will accompany each social organization format. These range from an "absolute" code of ethics for the calling to a "situation" ethic for the occupation. The crucial distinction is in terms of an absolute ethic versus an ethic which allows each individual to evaluate situations on their own merit to accomplish what the individual perceives to be the greatest good. Following Vietnam there has been a perceived loss of public trust and confidence in the military officer, indeed in public servants in general in the aftermath of Watergate and Koreagate. Futernick argues that public trust and confidence contribute to the intrinsic motivation characteristic of a calling and that their diminution creates an increased need for extrinsic motivators such as economic rewards. A vicious cycle thus has been created which contains the potential for the weakening of a professional ethic if the marketplace environment is seen as evoking a

situational ethic influenced by self-interest. If this state of affairs is characteristics of the older members of the officer corps, what are the implications for the socialization of military novices if the officer corps itself lacks a professional ethic and the prevailing social format of the organization is more conducive to the development of a situation ethic?

These are but a few of the possible implications for further study and questioning of contemporary military developments in light of Dr. Moskos' thesis. That such matters are of more than academic concern is evidenced by the fact that next week a wide array of Army personnel are to meet near Washington, D.C. to focus upon these issues and to determine how they affect the Army's command climate.

PANEL DISCUSSION:
The Expanding Role of Women in the Military*

Historically, with a few notable exceptions, women in the military have filled a few selected positions which have been traditionally recognized as "women's" jobs in the civilian sector. In recent years, there has been an increasing trend to utilize women, both officer and enlisted, in job roles which have traditionally been filled only by men. The members of this panel, representing the three branches of the military, are all deeply involved in the planning and implementation of programs to increase the utilization of women in a wide spectrum of career fields.

Co-Chair: Lt Col Eugene H. Galluscio and Capt David C. Gillman, Department of Behavioral Sciences and Leadership, USAF Academy, Colorado.

Discussant: Lt Col Verna Kellogg, USAF
Coordinator for Women in the Services
Washington, D.C.

Members: Colonel Edith M. Hinton, USA
Deputy Director Womens Army Corps (HQDA)
Washington, D.C.

Lt Col Shirley Bach, USAF
Special Assistant to the Secretary USAF
Manpower and Installation
Washington, D.C.

Lt Cmdr Joyce Killmer, USN
Head of Women's Equal Opportunity
Naval Personnel
Washington, D.C.

*Manuscript note available for publication.

THE UTILIZATION OF WOMENT IN COMBAT: THE ARMED FORCES OF GREAT BRITIAN, WORLD WAR I AND II

Nancy Goldman
University of Chicago

It is the usual strategy of social research to develop a number of case studies and then to pursue a more systematic and statistical analysis of a larger sample. But the reverse approach can on occasion be particularly revealing, and such is the logic of this paper. In my earlier paper, "The Utilization of Women in the Armed Forces of Industrialized Nations," I investigated 22 countries (Goldman 1977). From my sample the level of industrialization was unrelated to the utilization of women in a nation's armed forces. (It should be noted that in the developing countries women are almost totally absent.) But I found that internal democratic institutional practices facilitated the incorporation of women in a nation's military formation.

As a follow up, in this paper, I chose to investigate one case study in depth, namely Great Britain, in order to examine in detail the process by which women were introduced into the armed forces. The time period encompassed World War I and World War II and this case is important since the trend included the use of women in combat type (sometimes termed "operational") roles.

The question of women's utilization in combat has been raised in the 1970s and often the discussion and analysis proceed as if the issue were being faced for the first time. As a component of a larger study of the utilization of women in combat, sponsored by the Army Research Institute for the Behavioral and Social Sciences, the case of Great Britain during World War I and World War II is particularly relevant. Detailed biographical and archival records supply important sociological, psychological and organizational data on the problems and issues involved in the utilization of women in combat type roles and in combat units in Great Britain. This paper focuses especially on these conceptions.

Despite military policy which held that the women's auxiliary services of the British armed forces were non-combatant, in both World War I and World War II members of these auxiliaries were involved in combat. The tradition of women in the military, even before World War I was nursing or transporting the wounded at the front line, as well as behind the lines.[2] With the development of air power, often the posts of women behind the lines became target areas for air raids. In World War II, women in air defense combat operated anti-aircraft weapons and balloons. Another type of "weapon" utilized by military women was radar which enabled them to serve as combatants from the operations rooms in which they were posted.

Great Britain during World War I organized women's auxiliaries of the armed forces setting a precedent for women's utilization in the military on a far broader scale in World War II. While approximately 100,000 women were in service ibetween 1917 and 1919 in five categories of work, there were 453,200 in 1944 who, since 1941 had been employed in approximately 63 trades in the ATS, 65 trades in the WAAF, and the WR NS, whose work fell into four categories, were involved in 60 to 100 kinds of shore duty (International Labour Office 1964, pp. 140-41). The experiment of women's auxiliaries proved valuable to the military during World War I and conditioned public opinion as well. The traditional Victorian values were in flux and by 1914, women in the labor market were becoming common. The suffrage movement was in full swing with suffragettes demanding women's rights to the vote and full citizenship. London was a mecca of women's organizations and the war provided upper and upper-middle class, emancipated, women with a new cause around which to organize women's committees. Many paramilitary associations were formed as well as groups mobilized by the Red Cross, followed by the formation of the Women's Army Auxiliary Corps (WAAC). The derogatory attitude of the public towards women in the various organizations began to change once the Women's Auxiliaries became visible.

> "We in England grew so tired in the early days of the war, of the fancy uniforms that burst out upon women. Every other girl one met had an attack of khakiitis, was spotted as the pard with badges and and striped as the zebra. . . . With the official recognition of bodies such as the VADs and the even more epocmaking official founding of the WAACs the point of view of the ununiformed changed. The thing was no longer a game at which womon were making silly asses of themselves and pretending to be men; it had become regular, ordered, disciplined and worthy of respect. In short, uniform was no longer fancy dress." (Jesse 1918, p. 8).

Society had become accustomed to women associated with the Armed Forces nursing and or transporting the wounded at the fronts in Belgium, France and Siberia.[3]

With technological developments after World War I, the distinction between combat, combat support and non-combat roles became blurred, and the civilian kinds of duties which women performed for the military in World War I were expanded in World War II to include duties traditionally performed by men including semi-operational and operational jobs on anti-aircraft batteries, searchlight batteries, balloon barrages, and in Operations Rooms of Fighter Stations. The women's Auxiliary Air Force (WAAF) initiated work in radio-location (radar). The utilization of women in such non-traditional roles, along with the unprecedented conscription of women in Great Britain were difficult for the public to learn to accept. Rumors of immorality and drunkenness were widespread and it became necessary for a Committee on Amenities and Welfare Conditions in the Three Women's Services to be appointed by the government to investigate the

---

[1]Sponsored by the U.S. Army Research Institute for the Behavioral and Social Sciences. The views and conclusions contained in this document are those of the author and do not necessarily represent those of the Sponsor or the U.S. Government.

[2]Queen Alexandra's Imperial Military Nursing Service, organized in 1902; 1906 Territorial Force Nursing Service established; 1909 Red Cross and Order of St. John at request of Army Council, organized Voluntary Aid Detachments (VADs) to supplement the Territorial Medical Service. 1909 First Aid Nursing Yeomanry (FANY) formed as a mounted corps to transport wounded from battle field to hospitals.

[3]There were 45 nurses including VADs who are officially recorded as having been killed or drowned through enemy action (The War Office 1922, p. 194).

situation and allay fears and misconceptions of the public (1942).[4] (Such rumors had occurred during the first World War as well, and were found by an investigative committee to be largely false.) (Gwynne-Vaughan 1942, pp. 50-51). The military justified its reliance on large numbers of women in such a wide variety of roles as a product of the emergency situation. With the exception of direct, offensive combat roles, women were found in almost every position. Their air defense combat roles included service on anti-aircraft batteries as artillery women who operated the fire control instruments of the guns, determining whether or not the target was hit, raising and lowering the balloons, and manning plotting rooms located on fighter stations where they plotted the course of every ship and enemy aircraft and transmitted this information to anti-aircraft batteries.

Trends in Utilization

The army was slow to call on women in World War I to serve with the armed forces. In March 1917 the WAACs joined the VAD and FANY convoy drivers who were already in all the rear areas and base camps on the continent. The French and Belgian Red Cross had employed them as drivers and nurses and clerks since the beginning of the war. As early as 1914 there were a total of 47,196 VADs in the service (War Office 1922, p. 193). Preceding the official formation of the WAACs, members of the Cookery Section of the Women's Legion were employed by the military, and in February 1917 authority was extended to the Motor Transport Section of the Women's Legion (F.A.N.Y). When it was recognized by the War Council that 12,000 women could release men on the lines of communication for service at the front, the decision to organize the Women's Army Auxilary Corps was made and a few hundred were authorized to serve at home. On April 9, 1918, following the WAACs service while under shellfire, the name was changed to Queen Mary's Women's Army Auxiliary Corps (The War Office 1922, p. 205). The Navy organized the Women's Royal Naval Service (WRNS) in November 1917, and the Air Force organized the Women's Royal Air Force (WRAF) in April 1918.

The experiment of the women's auxiliary services was a success, as indicated by the increase in the number of women enrolled. The WAAC, with 2,377 women in August 1917 reached a peak of 40,850 women (or 1 percent of the army) in November 1918 (War Office 1922, p. 206). Comparative figures for the WRAF and WRNS are not available but it has been stated that between 1917 and 1919 approximately 32,000 had served with the WRAF and between 5,000 and 6,000 with the WRNS (Markham et al, 1942, pp.3-4). Although these services were formed on a temporary emergency basis and were dissloved with the termination of the War, they had laid the foundation for the newly formed services of World War II.

In 1938, with the threat of impending war the Auxiliary Territorial Service (A.T.S.) was formed by Royal Warrant. This organization incorporated the Emergency Service which had been formed in 1936 by some women officers of World War I in order to prepare women to serve as officers of any women's service of the armed forces in the event of war. The A.T.S. initially replaced the Queen Mary's Army Auxiliary Corps and the Women's Royal Air Force of World War I. When the Royal Air Force became a separate service it organized at the same time a women's division, called Women's Auxiliary Air Force (WAFF). The following year the WRNS of World War I was revitalized. Like their predecessors, these organizations were civilian, the women "camp followers," until April 25, 1941 when, as a result of the need to utilize women in combatant roles (Markham et al,1942, p. 4) the A.T.S. and WAAF were granted complete military status by the Government under the Order-in-Council, The Emergency Powers Defence (Women's Forces) Regulations (Secretary of State for War 1941, p. 1699). The Admirality chose to maintain the civilian status of the WRNS, saying that it was no less of an integral part of the Navy as a result (Markham et al 1942, p. 11). With full military status for the ATS and WAAF the military was able to utilize them in a wider variety of roles involving the handling of weapons.

> "The employment of ATS personnel to operate some of the many instruments used on anti-aircraft gun-sites had lately been recommended but could not be authorized under the 'non-combatant' definition in the Royal Warrant of 1938." (Gwynne-Vaughan 1942, p. 135).

In December 1941, under the continuing pressure of manpower shortages, and in order to qualify the combatant aspects of women's roles, the National Service (No. 2) Act was passed making women as well as men liable for conscription, and it provided that women would not be required to use lethal weapons unless they volunteered to do so—in writing (Markham et al 1942, p. 4). This rule was in practice since Spring 1940 when the initial experiment in training ATS members on anti-aircraft batteries was made.

The data on the utilization of women in the armed forces, as indicated in Table 1, show the tremendous growth of the services and the higher proportion of women utilized by the Air Force than by the other services. In 1939 there were a total of 36,100 women in the three services representing 2.27 percent of the armed forces while in 1943, 453,200 women represented 9.39 percent. Even before conscription increased the numbers involved (conscription was not put into effect for women until the end of 1942) (International Labour Office 1946, p. 140), by December 1941 there were already 205,100 women, or 5.39 percent of the armed forces were women. The Air Force, with the largest number of "non-combatant" roles, consistently had the highest proportion of women, although it was a smaller service in numbers of women than the ATS (with the exception of 1941, when there were 13,300 more women in the WAAF than in the ATS as a result of the unpopularity of the ATS at the time due to problems of disorganization and rumors). That year the 98,400 women in the WAAF represented 10.80 percent of the total Air Force, while for the ATS, the 85,100 women were 3.51 percent of the Army and for the WRNS, 21,600 women represented 4.59 percent of the Navy. The Air Force reached its peak in utilization of women in March 1943 when 180,100 WAAF constituted 15.96 percent of the Air Force, the highest proportion of women in the services to date. The comparative peak percentages for the other two services are for December 1944 when the WRNS reached 8.6 percent of the Navy, and September 1943 when the ATS was 7.35 percent. The WRNS was a much smaller service and more selective. Unlike the other two organizations, it accepted women only from medical grade 1 whereas the ATS and WAAF took them from grades 1 and 2.

[4]The *Report of the Committee on Amenities and Welfare Conditions in the Three Women's Services* was presented to Parliament in August 1942 (cmd. 6384). The appointment of the Committee was announced in the House of Commons on February 24, 1942 by the Prime Minister.

TABLE 1

NUMBER AND PERCENT OF WOMEN UTILIZED BY THE ARMED FORCES OF GREAT BRITAIN: 1939-1945
(by service and year)

Date	No. WRNS	% Navy	No. ATS	% Army	No. WAAF	% Air Force	Total in 3 Services	% in Services
Dec. 1939	3,400	1.56	23,900	2.08	8,800	3.93	36,100	2.27
Dec. 1940	10,000	2.92	36,400	1.72	20,500	4.01	66,900	2.26
Dec. 1941	21,600	4.59	85,100	3.51	98,400	10.80	205,100	5.39
Dec. 1942	39,300	6.49	180,700	6.58	166,000	15.06	386,000	8.67
Mar. 1943	45,000	6.87	195,300	6.92	180,100	15.96	420,400	9.13
June 1943	53,300	7.47	210,300	7.29	181,600	15.81	445,200	9.38
Sept. 1943	60,400	7.84	212,500	7.35	180,300	15.51	453,200	9.39
Dec. 1943	64,800	7.89	207,500	7.19	176,800	15.04	449,100	9.19
Mar. 1944	68,600	8.20	206,200	7.14	175,700	14.94	450,500	9.20
Dec. 1944	73,400	8.60	196,400	6.64	166,200	14.44	436,000	8.79
June 1945	72,000	8.42	190,800	6.13	153,000	13.87	415,800	8.20

SOURCE: Great Britain Central Statistical Office (1951: Table 10).

Sex Roles and the Integration of Women into the Armed Forces

Although established by the armed forces, the women's services in World War I were civilian and were adjuncts to the armed forces rather than integral parts of them. Women enrolled, they did not enlist; they had their own ranks which paralleled those of the military; officers of the WAAC were called "officials"; the army did not permit women commissions, or even use of the title "officer." Military law applied to women as civilians only where there was a serious violation injurious to the army; other minor offenses were handled in the civil courts. Without military status women in World War I were limited to administrative, domestic, transport and other "traditional" women's roles. These roles, however, provided women with the armed forces an opportunity to prove themselves capable of replacing men behind the lines and capable of serving under enemy attack. This involvement of women in air attacks to some extent implemented their acceptance into the military. As Dame Helen Gwynne-Vaughan, first Chief Controller of the WAAC, later Commandant of the WRAF and finally Chief Controller of the ATS, said:

> "Had it been considered necessary to remove them at every hint of danger, thus disorganizing the officers, our value to the army would have disappeared." (Gwynne-Vaughan 1942, p. 55).

To be involved in dangerous events was an experience valued by the women since it signified their participation as soldiers. When an aerial torpedo exploded in a trench killing nine and wounding six others of a group of newly arrived WAACs,,those who had preceded them felt it was inappropriate that newcomers had this "honor" (Gwynne-Vaughan 1942, 0. 57). It was the first occasion when women were given a military funeral, and military medals were awarded to four women who had performed heroically during the crisis. This event gave the WAACs visibility and enhanced their status as "soldiers." On an occasion when an officer, in response to a request to honor the outstanding conduct of the women during a raid by mentioning them in the Orders, said: "We do not thank soldiers for devotion to duty; we do not propose to treat women differently," they felt they had been admitted to a great fellowship. [5] And when they were told that it was expected that some women would be killed they "shared the sense of superiority" in sharing this knowledge, of their shared risks of danger.

[5]In World War II, there were 1,486 casualties of members of the women's auxiliary services recorded: 624 killed, 98 missing, 744 wounded and 20 prisoners of war. Of those who died, 124 were WRNS, 751 ATS and 611 WAAF. (Central Statistical Office 1951, p. 13).

Women in World War II auxiliaries were exposed to the same dangers of air attacks but to a far greater extent, and there are many reports of heroism and many more military medals were awarded. Fighter stations and ports where women were posted working with radar, signals equipment and other instruments were frequently raided. Nevertheless, the women were reported to have continued fulfilling their responsibilities—even under the worst situations—for which they earned respect from their male co-workers—as well as from civilian society—and greater self-esteem in general.

The expansion of women's roles into the semi-operational and operational aspects of the armed forces in World War II served as an important means of integrating them into the military establishment. Many of the limitations previously imposed upon women's potential roles were removed with the designation of their full military status.

> "So often the fact that we were civilian had blocked some avenue of usefulness or made impracticable for our women some benefit which existed for soldiers. In particular the employment of ATS personnel to operate some of the many instruments used on the anti-aircraft gun-sites had lately been recommended but could not be authorized under the 'non-combatant' definition of the Royal Warrant of 1938." (Gwynne-Vaughan 1942, p. 135).

Women were serving on gun-sites, aerdromes, balloon barrages, as coders, signallers, plotters, in radio-location, handling of projectors and range and height finders. The searchlights in the home establishment were almost entirely taken over by the ATS. Many women had shared training experiences and combat experiences with the men, having been trained in coeducational groups and serving in mixed units. The women felt themselves an integral part of the military. The Committee on Amenities (1942, p. 9) reported:

> "We do not find that hardships or rough conditions were resented on the operational sites; but on the contrary, they were willingly accepted by women who feel, when serving guns and balloons and radio-location, that they come near the actual war and that they are playing a part in the struggle."

Coeducational training aided in the integration of women into mixed units. At the practice camps girls on radar and other instruments worked together with male gunners so that when they left the camp for the gun-site they were accustomed to the arrangement. In addition, an effort was made to select young male recruits, who were used to coeducation in their civilian life, to work in mixed units rather than old soldiers who were conditioned to an all male military.

Although women played a vital part in air defense by means of anti-aircraft guns, balloons and radar, the women's services were considered non-combatant and women were barred from service at the front line, (even drivers for the ATS were not allowed at the front, as they had been earlier when they were members of independent societies such as the Mechanized Transport Corps). (Leslie 1948, p. 9). They were barred from service aboard aircraft (except for test pilots and nurses who flew to the front line in order to return with the wounded)[6] and they were barred from ships at sea. Women on anti-aircraft gun-sites operated instruments involved in controlling the fire of guns, but they were not allowed to fire them. One excuse given for this limit set on women's roles was that the War Office was concerned that the Germans might treat women as combatants if they had any training in "musketry" (House of Commons Debates 1943, p. 2180). Another was that the shells were considered too heavy for women to load and unload (Wadge 1947, p. 120). Women were in complete charge of many balloon barrages despite the very strenuous work involved in operating the balloons which were so effective in destroying flying bombs. Women also "fought" in air and sea battles from plotting rooms in the operations headquarters of air and naval stations where they plotted the course of allied and enemy ships to assure the accuracy of the bombing.

Conclusion

Once women were trained in skills useful to the military and once they requested to be allowed to join its forces, they were called on to serve as an emergency measure to meet manpower shortages, first during World War I and again during World War II at which times they formed women's auxiliaries of the armed forces. Prior to World War I women served in the military as nurses, frequently tending the wounded at or near the front.[7]

Industrialization which mechanized war and created new manpower needs also taught women to work outside the home in occupations which were useful to the military as well as to the civilian sector. Therefore, with the outbreak of total war, Great Britain with its manpower shortages and as a democratic country which wished to allow women to volunteer to participate in all aspects of the war effort, turned to women and organized women's auxiliaries to serve, initially, the lines of communication in France, in roles for which they had been trained in civilian life. In World War II, with technological developments which affected military operations, greater personnel needs were created and Britain called on a large number of women to perform a wide variety of roles, which included most aspects of air defense.[8] Consequently, once granted military status, women were eligible for and became involved in combat type activities, despite military policy which provided that they were a "non-combatant" force.

There is much to be learned from the assessment of the utilization of women in combat in Great Britain. It is clear that technological trends, and especially "high technology" facilitate the incorporation of women into the armed forces, and into combat type roles. But it is also clear that we are not dealing with the formulation and adoption of explicit military personnel policy. Rather we are dealing with a gradual process of change, with incremental developments which produce operative modifications even though the older language and symbolism remain formally in effect. We are dealing with an element of ambiguity. This is the process in which the United States appears to be engaged as it implements the all volunteer force during an historical period in which there is strong societal emphasis on women's equal rights.

---

[6]Women of the Air Transport Auxiliary, which was a civilian organization under the Minister of Production and was never militarized, ferried every type of aircraft including four engine bombers.

[7]There were many incidents reported of military nurses during World War I and II serving "near" the front. For example, British nurses with the British Army of Liberation worked in Normandy in hospitals which were so close to the front line that shelling and bombing were a common occurrence within a half-mile of the hospital. The nurses wore battledress and lived in the same style as the troops. See British Official Photograph in Wadge 194, p. 28.

[8]The higher number of personnel in service during World War II than in World War I was mostly the result of the mobilization of women. The number of men in 1918 was approximately that of 1944 (The Prime Minister 1944, pp. 4, 5, 8).

REFERENCES

Beauman, Katharine Bentley. 1943. *Wings on Her Shoulders*. London: Hutchinson.

Douie, Vera. 1949. *Daughters of Britain*. Oxford: George Ronald.

Goldman, Nancy L. 1977. "The Utilization of Women in Industrialized Nations." *Sociological Symposium* 18 (Spring): 1-23.

Great Britain Central Statistical Office. 1951. *Stationery Office and Longman's Green.*

Gwynne-Vaughan, Helen. 1942. Service with the Army. London: Hutchinson.

International Labour Office. 1946. *The War and Women's Employment*. Montreal: International Labour Office.

Jesse, F. Tennyson. 1919. *The Sword of Deborah: First Hand Impressions of the British Women's Army in France*. London: Heinemann.

King, Alison. 1956. *The Story of Some of the Women Ferry Pilots of the Air Transport Auxiliary*. London: Pearson.

Leslie, Anita. 1948. *Train to Nowhere*. London: Hutchinson.

Markham, V., et al. 1942. Cmd. 6384. *Report of the Committee on Amenities and Welfare Conditions in the Three Women's Services.* London: His Majesty's Stationery Office.

Ministry of Information. 1943. *Roof Over Briatin: The Official Story of Britain's Anti-aircraft Defences, 1939-1942*. London: His Majesty's Staionery Office.

Mitchell, David. 1966. *Women on the Warpath*. London: Jonathan Cape.

The Prime Minister. 1944. Cmd. 6564. "Statistics Relating to the War Effort of the United Kingdon." In *Parliamentary Papers* 4 Ses. 24: 8.

The Secretary of State for War. 1941. "Auxiliary Territorial Service and Women's Auxiliary Air Force (New Status)." In *House of Commons Debates*. 5s. 370.

Wadge, D. Collett (ed.) 1947. *Women in Uniform*. London S. Low Marston.

The War Office. 1922. *Statistics of the Military Effort of the British Empire during the Great War, 1914-1920*. London: His Majesty's Stationery Office.

Ward, Dame Irene. 1955. *FANY INVICTA*. London: Hutchinson.

Whateley, Dame Leslie. 1949. *As Thoughts Survive*. London: Hutchinson.

"Women in the National Service." 1943. In *House of Commons Debates*. 5s. 391.

ARMY ENLISTED SEX ROLE ATTITUDES: A MALE-FEMALE COMPARISON

Jack M. Hicks
Army Research Institute For The Behavioral Sciences

The present investigation grew out of a need of the U.S. Army to etter understand the contributions which young women might potentially make to its manpower system. As a first step, this effort was addressed specifically toward generating a better understanding of the perceived status and utilization of enlisted women in the Army. This objective was implemented primarily by comparisons of enlisted women (EW) to enlisted men (EM) on a variety of attitudinal and perceptual dimensions.

Based upon the results of focused group interviews, and other hypotheses, a survey questionnaire was designed, and pilot tested in November, 1975 at Fort Ord, California. The finalized questionnaire consisted of separate, though similar EM and EW versions.

Nine Army installations with a high percentage of EW were selected for data collection. These posts yielded a total of 1,718 EW, and 835 EM respondents. To the extent possible, selection of participants was governed by the following specifications: length of service 18-24 months (first term); primarily E-3 and E-4; eligible for reenlistment; and diversity of MOS.

The questionnaire was administered to sex-segregated collectives of 25-100 respondents per session, in December, 1975.

One question of interest concerned opinions about which jobs were best suited to which sex. In general, there was remarkably high agreement between the sexes in this regard. The best illustration concerned those jobs into which women were not currently permitted entry. The respondents not only generally agreed with the opposite sex, but also among themselves that the fighting and heavy equipment MOS's were best suited to men. This consensus was particularly striking with regard to combat, where only one fifth of each group considered combat suitable for both sexes; with fully two thirds of each group agreeing that combat was for men only. Also, only slightly over one third of *both sexes felt that "men and women can fight side-by-side in the battlefield".*

Of related interest was "how much would you like to go into combat" in case of another war. Twenty-nine percent of the EW responded positively. Though men, in general, showed a greater inclination than women, with 40% favorable, a sex comparison must take into account the EM already in combat MOS's.

General agreement between and within sexes was reached concerning jobs in which women have *traditionally* functioned in the Army. The overwhelming consensus of both sexes was that these jobs were suitable for either sex. Also, there was very little sentiment by either sex favoring the suitability of *any* Army MOS for "women only". The area in which sharpest disagreement between the sexes was found, also was the area of greatest division of opinion *within* the EM and EW groups. This concerned jobs in which women had *not* traditionally been permitted entry, but *which had recently been opened to women.* In these areas, there was a much greater tendency for EM than EW to feel that such jobs should be done by *men only.* Even so, men were sharply divided among themselves, with a slight leaning in favor of regarding these areas as appropriate for both sexes. This was particularly true of military police, which over half of the EM reported to be suitable to both sexes, with only 23% claiming this MOS for men only.

With regard to attitudes toward the "overall quality of men in the Army", the EW tended to hold the EM in modest esteem. A significant plurality (41%) of women rated the men as "average", with only 19% giving above average ratings. The women fared somewhat better in the eyes of the men, with 34% of the men giving average ratings, and 28% above average. These ratings were not appreciably different from those that the men gave to the men. The women were not so generous, however, rating the women generally much higher than *they rated the men.* One factor which may have influenced such incomplete reciprocity related to higher mean aptitude and educational levels of EW. However, unwillingness of the men to rate women higher than men may relate to their perceptions of compensating attributes on the part of men--such as physical strength, outside skills, and the like. If so, it was evident that the EW did not fully share such perceptions. In fact, an overwhelming majority (82%) of the EW, as compared to 59% of the EM, agreed that "in some ways, women in the Army are sharper and better soldiers than a lot of men.."

This was not to suggest that the EW personally identified as soldiers more so than the men. Upon being requested to choose those adjectives which best described various concepts, the EW did not depict themselves as approximating the attributes of the "ideal soldier" to the extent that the men did. The EM attributed the ideal soldier such primary designations as honest, level-headed, friendly, helpful, and straightforward; precisely the same set of adjectives with which they described themselves. The women were in remarkable agreement with the men as to the primary attributes of the ideal soldier. However, they deviated significantly in describing themselves, tending toward a combination of the ideal soldier and the ideal woman. Though the EW self-image retained much of the ideal soldier, there was also a tendency to share such attributes of the ideal woman as femininity, warmth, emotionality, and even stubbornness, to a greater extent than they considered appropriate for the ideal soldier.

It was also of interest that both sexes tended toward clearly differentiated stereotypes of women in traditional as opposed to nontraditional job roles. The EM described the "traditional MOS woman" as warm, friendly, helpful, feminine, and soft; attributes resembling those the EM felt best described the ideal woman. By sharp contrast, the "non-traditional MOS woman" was described by the EM as pushy, masculine, troublesome, unattractive, and hard. The EW, as a group, described the nontraditional MOS woman in similar socially undesirable terms. Though women actually in nontraditional MOS's were somewhat more lenient in their descriptions of the nontraditional MOS woman, even they were substantially divided and not as positive as might be expected. Only 60% of nontraditional MOS women expressed "a lot of respect for women who go into nontraditional MOS's".

Another finding was the considerably greater percentage of men than women who indicated that the sexes should be treated equally in the Army. Over three fourths of the EM and only 62% of the EW seemed to partake of the egalitarian point of view. There were indications that many of the men felt that there was partiality toward women in the Army, and that their strong indorsement of equality of the sexes was actually a plea for *equality for men* rather than women. For example, 62% of the EM agreed that a "lot of women. . .use their sex appeal. . .to get promoted or get special favors and jobs". Over half of the EW also agreed with this statement. Also, 54% of the EM, as compared to 31% of the EW, agreed that "a woman in the Army will use her femininity to get away with as much as she can".

Returning to combat attitudes, it will be recalled that women were attributed far lower combat credibility than men, by both men and women. It should be added, however, that three fourths of both sexes also agreed that "women in the Army should be trained in handling weapons". Moreover, a clear minority of both the EM (27%) and the EW (21%) felt that "a woman isn't capable of fighting in a war." There was, however, some disagreement between the sexes regarding combat-related MOS's and combat behavior. Sixty-one percent of

the EW, as compared to 47% of the EM, felt that "women should be allowed into combat-related MOS's it they chose". Thirty-one percent of the EM and only 18% of the EW agreed that "in the battlefield, a woman would panic and run under fire". Though these differences were real, those of both sexes who showed strong disrepect for the combat-related capabilities of women were clearly in the minority. Thus, the overall sentiment pertaining to the female in the combat environment was one of essential compatibility between the sexes across several levels of participation. Greatest accord was reached with regard to the prospect of full participation of women in combat (which both sexes rejected), or merely training women in handling weapons (which both sexes indorsed). Somewhat less agreement was reached at the intermediate level, namely, the combat-related MOS. Both sexes, particularly EW, stopped short of suggesting that women were not capable of fighting, even if fired upon.

In conclusion, it can unequivocally be stated that most of the EM participating in the investigation felt that women can and do make a substantial contribution to the Army. Less than one fifth of the EM felt that "the Army is a man's world and women can't really belong in it". However, though the women were generally viewed as making a contribution, they were also viewed as *limited* in the contributions they could make. Areas in which the men least appreciated the potential of women concerned combat, combat-related, and certain other nontraditional female job roles. Such a tentative indorsement of women in the Army also seemed to be manifested in a kind of conflict as to how women should be treated by the men. There emerged a pattern that women should be both protected from the rigors of heavy duty, and be forced to carry their load. Sixty-four percent of the men felt that men should help the women with heavy lifting, but 76% felt that the sexes should be treated "equally".

Sixty-nine percent of the women also felt that the men should help the women with heavy lifting. The majority of women were also in agreement with the majority of the men that women should stay out of combat--at least alongside men. There was less agreement between the sexes regarding female participation in combat-related, and nontraditional activities generally. This issue appeared in the form of a perceived second-class citizenship among the women. In the final analysis, it was clear that the women generally felt that they could and should participate more pervasively (though not universally) in the military process.

UTILIZATION OF WOMEN IN NON-TRADITIONAL MILITARY CAREER FIELDS

Cecile S. Landrum
Analysis & Evaluation Group
ACS Studies & Analysis, HQ USAF
Washington DC 20330

Abstract

Women are entering the labor force and particularly the military in greater numbers than ever before. Policy changes concerning marital and family status and career opportunities have found women in jobs never before envisioned.

Consistent with all working women, military women encounter problems which negatively impact on their ability or on their performance. The basis of many of these problems are attitudes which are based on emotion rather than logic and which require motivation and discipline to effect positive change.

This paper will examine the impact of these pervasive attitudes on Air Force women as they enter a growing number of career fields in numbers that far exceed any previously anticipated.

Today's Women and the All Volunteer Force (AVF)

The role of women in the military is no longer to merely fill gaps in male recruiting and is in the forefront of All-Volunteer-Force considerations. A predicted dwindling manpower pool of eligible young men and an all-time high first-term attrition rate have led us to consider increasing the number of women in the Service and may demand a national policy response.

Recent Pentagon plans reveal an intention to increase the number of enlisted from 104,000 to 199,000 over the next five years. More significantly, women now serve in such non-traditional fields as electrical equipment repair, aircraft maintenance, forklift and truck operation, aircraft pilot operation, communications, and telephone line repair.

Policy decisions have changed concerning the kinds of jobs women can perform, the number of women who can serve, the locations in which they can serve, and their status. But increasing the number of women will not alone resolve the problems of the AVF. For women to succeed, their capabilities and desires to serve must be recognized and supported. Otherwise, the number and quality problems of sustaining the AVF could be replaced by problems of managing a military with significant joint-spouses and family concerns, or perhaps even by significant attrition problems occurring at different career points than being experienced now.

The Dilemma of Combat

Combat issues rather than full-utilization-of-women continue to surface and generate emotional arguments. The real issue, however, is that existing combat laws preclude women from full participation, although they are demonstrating by their high performance standards that they want to serve as equal partners.

Recent suggestions to open temporarily some combat-related duties raise questions about required manning in a combat environment. Our country cannot afford to gamble with its mission readiness, nor with the future of those women who have volunteered to serve their country. Sending women into combat assignments on a temporary basis should be totally unacceptable to any American who places a high priority on peace.

The Enlisted Women Today

The major numerical increase of women in the last five years has been in the enlisted grades. Changes in policies and legislation concerning marriage, pregnancy, benefits, ceilings, and promotions have served as catalysts for recruiting these increased numbers of women. But this rapid increase and the orderly and systematic personnel promotion system have created a situation in which 85 to 90 plus percent of the women are clustered in the lower four grades. The Services therefore lack a substantial number of women leaders and role-models in both the officer and upper enlisted ranks to support these young women in their new endeavors. During a recent field trip by the author, many women serving in non-traditional jobs expressed frustration about their lack of women supervisors. They were also concerned about the possibility of moving into desk jobs as they moved up in grade.

Many of the enlisted women acquiesced to accepting non-traditional jobs through the guaranteed enlisted program, but found themselves confused when they realized the full scope of the jobs. They got caught in the struggle for full opportunity and equal responsibility on the one hand, and their desires to be properly utilized on the other hand. As women enter the Services with more prior exposure to and understanding of non-traditional jobs, as neutral mental and physical job standards for all jobs are established, and as the women grow in numbers in the Service—these struggles can be resolved.

Women in the Armed Forces—Attitudes

Pervasive and complex attitudinal issues have in the past played a subtle but significant role in relationship to all other issues dealing with the utilization of women. Attitudes of man and women, as well as institutional attitudes clearly impact on the ability of women to perform their jobs. Women have been at a real disadvantage because of the lack of institutional and social cohesiveness. Where the law failed to inhibit women, policies—written and traditional—have continued to preclude them from substantial job roles. But as practices change, attitudes change.

Male Attitudes

Although the incidence is subsiding, some men still fear that the presence of qualified women in non-traditional jobs undermines their position of dominance and threatens the image of their being hard-fighting men in a chosen male environment.

There are still men who are paternalistic toward women, and some men who are resentful of women. Some feel that women should be placed on a pedestal and should not be exposed to distasteful tasks. During a recent visit to Israel the author found a protective attitude toward women to be prevalent in their cultural attitudes despite the ever present threat of war. In our society, however, such a protective attitude would preclude women from doing their fair share and their male co-workers could feel imposed upon. There are men, however, who feel that women shirk their full responsibilities by desire or policy and remain resentful of the women, believing that their presence results in added physical labor, longer rotations away from home, more remote assignments, late night shifts, and more doubling-up on workloads for the men in the unit.

Many men admit that they first view women as sexual objects—and only later see them as professionals. Some can never separate the two. While they want women to be feminine, many men find themselves uncomfortable working with women as their peers. Some find that once they work with women, their preconceived ideas change and become positive—especially when they know the women are trying their best; others fail to recognize that they are unable to communicate dissatisfaction and may be dissatisfied, no matter what.

Men also vary in their concepts of equal treatment of women. Some feel that equal treatment means treating and expecting women to act like men. Others think it is keeping women in traditionally female jobs. This non-standard treatment causes women to become frustrated, and they react to men's reactions, while attempting to maximize their performance.

Women's Attitudes

When women first moved into the "new" world of non-traditional jobs, many acted in traditional feminine ways as their dominant *modus operandi*, probably because it was their only known successful behavior in a male environment, not realizing that such behavior could polarize the situation. On the other hand, "superwoman" roles were assumed by many women to achieve a level of success, as they still find that they often have to work twice as hard to be considered half as competent as their male counterparts. Failure is often blamed on sex differences rather than on inability to perform. It has been difficult for women to learn to turn away from dual standards and to accept responsibilities that are not always attractive in order to achieve real equality. As more women move up in grade, they will provide more role-models in the non-traditional jobs.

Socialization of Women

Consideration of the roles of women must also address the psychological and educational preparation to fill those roles. New life-styles are still proving very threatening to many women. According to Matina S. Horner, President of Radcliffe College, women have historically "converged on the idea that feminity and individual achievements which reflect intellectual competence or leadership potential are desirable but mutually exclusive goals. The aggressive and, by implication, masculine qualities inherent in a capacity for mastering intellectual problems, attacking difficulties, and making final decisions are considered fundamentally antagnostic to or incompatible with femininity." Horner sees women going as far as purposely avoiding success in fear of negative social ramifications. Another study, implied that the best adjusted women were the underachievers who did not prepare themselves for an unrealistic future by overemphasizing academic performance.

From women's formative years, the sense of individual competition has been ingrained and has carried into the market place as well as in their personal lives. As women moved into the traditional male domain, a "queen bee" syndrome ensued as the result of tokenism. One or two successful women in an organization diminished the urgency to recruit more competent women, thus slowing down integration. This is changing as women are entering these career fields in greater density. For example, Yale University, upon opening its doors to women, found that until women accounted for 20% of the classes, they did not adjust or fare well.

Marriages

The number of Service marriages has risen as dramatically as the number of women entering the Service. Already over 40% of the Air Force's enlisted women and officers are married, and about 80% of the marriages are to other military members. When only spouse is in the Service, PCS's reflect personal desicions. Joint-spouse change of station for schooling and assignments conducive to career advancement that require separation become another matter.

Economists see substantial savings with one PCS move to fill two military vacancies. But a new set of assignment problems is created by moving increasing numbers of marrieds together and matching skills, training, and schooling assignments. Training and schooling, while critical to career enhancement, may place the assignments of the spouse out of sequence, cause the temporary overmanning of a position, move someone to accommodate a spouse, or incur additional moves to keep them together.

The Army has found that training exercises that include joint spouses with dependents are increasing, and already result in undermanning because one spouse is left behind to care for the children. Investments in potential leaders are lost when opportunities are denied because of the desire for joint assignments. This can be costly both economically and in terms of the numbers of quality people. As a result, a need is growing for realistic joint career planning which allows for career progression for both spouses, not just the wife tagging after the husband.

Pregenancy Policies

The increasing number of young military women and marriages has led to increases in pregnancies. While statistics indicate that Air Force women have half the fertility rate in comparison to the U.S. population and it has been determined that pregnancy is a "temporary condition,"—the system of movement creates a situation in which certain jobs are continually filled by young women, those who are most likely to become pregnant. Relegating women to desk jobs during this period can have negative impacts on both force readiness and career advancement because it can take the woman out of her occupation and may leave the job unattended.

While healthy women in traditional roles can continue their normal physical activity during pregnancy, problems relating to size, balance, and environmental hazards in non-traditional jobs are encountered. Pregnancies create temporary physical changes and often present emotional stress to women. An unmarried woman faced with pregnancy must further examine her ability to assume the demands of military life—especially when her job requires full-time commitments, such as with academic life, or more than normal field duty, or time away from home.

The highly structured military personnel system has also limited women's upward mobility in particular with respect to policies toward women who bear children. In the past, policy required military women to choose between career or family, usually at a time when career opporunities were about to develop for them, because pregnancy meant automatic discharge. Despite the policy changes in 1971 and 1975 relating to pregnancy and minor dependents, the military mission impacts on women in making decisions to manage both career and family. A choice that becomes increasingly difficult with increasing level of responsibility, a choice faced by young women in all career fields today.

The Impact on the Family

Single parents are often thought of in terms of being women only. However, there are far greater numbers of single fathers than single mothers in the Air Force today. This creates child care needs for all children, especially during deployment, exercises, shift night mission.

With more single heads of households, more homes with both parents husband situations, the changing complexion of the American family puts policymakers in the position of having to make plans with little empirical data. Sociologist Hamilton McCubbin, in recognition of the ensuing problems in recruiting and socializing high-quality military personnel, asks the following questions: "How will these changing situations affect military job assignments, family relocations, and extended separations? And will members of military families become less dependent on the system, more assertive of their personal and family needs, and less willing to subordinate their lives to the orders of the military establishment?"

Conclusion

Women still seek traditional jobs for pay and job security. The uncertainty in accepting additional responsibilities along with the assumptions of rights can be very intimidating. Some women, not ready for these responsibilities, capitalize on feminist traits and adopt double standards as a way of life.

In moving into the non-traditional jobs, women are moving into a male-oriented world. Changes in policies do not not necessarily mean changes in standards. But it is necessary for people to become aware of and sensitive to the prevailing male atmosphere in establishing job standards that reflect both mental and physical needs in order to avoid mismatches, unreal expectations, and eventually an aura of failure. The lines of command must be sharply defined from the start.

It is important that the interlocking issues described should by no means preclude women from fully serving in the military—especially at a time when their presence is becoming more important in making the AVF work. What is essential is that the women and men joining the military understand the mission and its priorities—and that the policymakers consider these issues in defining their requirements.

Whether these women remain in the military and contribute to our National defense is directly dependent upon the parameters within which they are allowed to serve—as equal partners or as an adjunct and disadvantaged minority. If this challenge is successfully met, the Defense Department will have been in the forefront of one of the major social accomplishments of this century.

REFERENCE NOTES

1. Utilization of women data from Department of Defense, Annual Report Fiscal Year 1979, p 328, and Office of the Asst Sec of Defense (Manpower, Reserve Affairs, and Logistics), "Use of Women in the Military," May 1977, p 7, Table 3.
2. Much of this paper will be based on recent field trips the author has taken to CONUS and overseas bases where she conducted on-site interviews with Air Force women serving in non-traditional jobs.
3. Much of the section of attitudes is based on the author's paper, "But Can She Type," which was presented at the Regional Conference of the Inter-University Seminar on Armed Forces and Society, Maxwell AFB, October 1976. Many of the items are substantiated by surveys. In particular, see Joel M. Sevell and Barry Collins "Soldiers' Attributions of Contemporary Vs Traditional Sex Role Attitudes to Themselves and Others," Research Memorandum 75-7 (U.S. Army Research Institute for the Behavioral and Social Sciences, July 1975, Kathleen Durning and Sandra Mumford, "Differential Perceptions of Organizational Climate Held by Navy Enlisted Women and Men," NPRDC TR 76 TQ-43, 1973.
4. Reference to Israeli Forces based on the author's visit to the Israeli Defense Forces, Feb 7-12, 1978.
5. Socialization of women studies have been addressed by Matina A. Horner, Journal of Social Issues, Vol 28, #2, 1972, p 158; and E.C. Lewis's "Developing Women's Potential," Ames, Iowa, Iowa State Press, 1968, in reference to an article by Ravena Nelson, "The Changing Image of the Career Women," (Berkely, Univ of California) p 35, concerning Vassar women.
6. Reference to integration of women at Yale University from personal interview with Jackie Mintz, Provost, Yale University, July 1975.
7. Reference to recent trends of college students. Personal interview with Matina Horner concerning college students choosing between careers and family, April 1977.
8. Data concerning enlisted women show that in 1964, enlisted women accounted for .9% of the Force—in 1977 they reached 6.7%.
9. Air Force statistics furnished by the Department of the Air Force, Military Personnel Center, Randolph AFB, Nov 1977; and Dept of the Air Force, Surgeon General letter, subject: "Non-effectiveness Rates by AFSC, Sex, Disease, Category and Age," 29 July 1976.
10. Reference to military families from Hamilton I McCubbin and Martha A. Marsden, "The Military Family and the Changing Military Profession," paper presented at the 1976 Regional Meeting of the Inter-University Seminar on Armed Forces and Society, Maxwell AFB, October 22-23 1976.

RESEARCH ON UTILIZATION OF WOMEN IN THE AIR FORCE

by

Thomas W. Watson
Jeffrey E. Kantor
Personnel Research Division
Air Force Human Resources Laboratory

The military environment is constantly subject to change resulting from external and internal forces. External influences impacting on the military include: reduction in the pool of potential recruits; recruiting difficulties in an all-volunteer environment; increased competition with academia and industry for personnel resources; and changes in sex roles and occupational aspirations, especially for women. Internal forces currently modifying the military include: reduced financial allocations, increased personnel costs, reduced personnel strength authorizations, increased emphasis on compliance with federal directives concerning equal employment opportunity, and changes in service policies such as the expansion of the number of women and the pattern of their utilization to include practially all occupational specialties. These influences have produced various managerial problems. The military has had to do more with less resources, and it has become crucial that each military position be filled with the best person for the job, whether male or female, and to ensure that incumbents are effectively used, relatively satisfied, and productive. In order to best serve Air Force needs and interests of personnel are well used, women are now entering non-traditional careers previously reserved for males. In January 1978, only seven specialties remained closed to women. However,by March 1978, this number decreased to three and may be further reduced if legal and policy changes concerning the combat role of women are adopted.

With the increased influx of women into the Air Force (from 40,000 in 1977 to a projected 81,000 in 1983) and the diversification of their utilization, a number of questions need to be answered concerning screening and selection procudures; classification and assignment; training, utilization, and performance; and attrition and retention. Specifically, can screening, selection, and assignment procedures be equally applied to, and be equally effective with, both sexes; do similar factors affect the training and job performance of both sexes; and do male female differences exist which should be considered when policy or personnel decisions are being made?

The Personnel Research Division of the Air Force Human Resources Laboratory is involved in a continuing research program designed to obtain information needed to answer questions of the sort mentioned. Some of this research has had a specifically female orientation. However, in a recent meeting of the AFHRL Technical Advisory Board, it was decided that future research should investigate female management problems within a broader, total-force context which includes simultaneous study of, and comparison with, males. Also, it should be noted that the research projects discussed here are, for the most part, not complete, and reported data represent only fragments of the anticipated products. Therefore, this presentation should be considered an overview of what is being done rather than a snyopsis of what has been accomplished.

Many of the individuals who leave the service prior to the normal termination of their first term have experienced problems in adapting to military life. Identification of these personnel before enlistment is a goal of screening, and a suitable screening instrument, applicable to both sexes, is needed. Previous research on a sample of male, draft-motivated enlistees (Guinn et al., 1975) has led to the conclusion that the History Opinion Inventory (HOI), a 100-item objective questionnaire, has potential usefulness as a screening device for suitability for military service. Current research is continuing the evaluation of this instrument on a 1975 sample of 32,000 male and 8,000 female subjects. Initial analyses indicate that the HOI is applicable to both males and females. Separate set-specific scales were more valid in predicting failure in basic military training than was a unisex scale. As technical training and job performance criteria become available, additional analyses will be accomplished.

A related project was initiated in 1977 at the Armed Forces Entrance and Examining Stations (AFEES). A shortened version of the HOI, the 50-item Military Screening Inventory (MSI), was administered to approximately 77,000 potential recruits from all services. AFHRL follow-up of these personnel will involve analyses to determine the overall utility of the MSI as well as its applicability and validity for males and females.

To increase the precision and efficiency of the classification and assignment process, information in addition to aptitude data is desirable. Therefore, the Personnel Research Division has been involved since the early 1970's in the development and validation of a vocational interest inventory designed for use with Air Force enlisted personnel. This instrument, called the Vocational Interest Career Examination (VOICE), has been validated for men and women against a criterion of job satisfaction, compared with similar instruments from other DoD departments, and normed on a high school sample. Also, a non-verbal version is in the early stages of development.

Alley and his associates (Alley et al., 1977) developed 18 homogeneous, factor-referenced subscales for the VOICE using an Air Force sample of 10,000 males and 12,700 females. In the development of these scales, inter-item correlations and factor structures were found to be virtually identical for males and females. While the subscales had the same meaning for both sexes, the degree of preference associated with each subscale differed by sex group. Males typically expressed greater interest in Electronics, Heavy Construction, Mechanics, Law Enforcement, and Marksman scales while females preferred Office Administration, Medical Service, Aesthetics, Food Service, Audiographics, Agriculture, and Teacher Counseling. A later study by Berger and Berger (1977) concerned with the norming and standardization of the VOICE in a nationwide high school sample confirmed these results, with the exception that females in their sample expressed less interest in Science than did males, while in the previous study no sex differences existed on this subscale.

Alley et al. (1976) sent a job satisfaction survey to a subset of approximately 18,000 of his original sample who had been on the job for between 8 and 12 months in order to obtain a criterion against which to validate the VOICE scores and expressed satisfaction by sex, aptitude index and occupational group was examined. Aptitude was found to contribute little to prediction, but satisfaction differed considerably by sex and DoD occupational group. Women were found to be experiencing difficulty in adjusting to assignment in some non-traditional areas. Relationships between prior interests and subsequent job satisfaction were found to differ between men and women assigned to general administration career areas with satisfaction among females being more highly related to prior interests than was satisfaction among males. Similar sex differences were noted for the general aircraft mechanic specialties with VOICE subscales being more predictive of satisfaction for females than for males. When regression weights were developed for each of the DoD occupational groups, it was discovered that, with the exception of one administrative and two mechanical occupational groups, identical weights could be used for predicting male and female job satisfaction.

An interesting spin-off of this research is the comparison of expressed job satisfaction between male and female respondents doing similar jobs. Females were found to prefer Medical Care; Miscellaneous Medical and Dental Specialties; Technical and Allied Specialties; and Material Receipt, Storage and Issue while males preferred Radio/Radar Equipment Repair, Radar and Air Traffic Control, Miscellaneous Communications and Intelligence Specialties, Aircraft Mechanic, and Service and Supply.

Although additional sex-differentiated VOICE studies are planned, such as validation against long-term satisfaction, technical school training and job performance as well as norming of VOICE instruments which have been scored using alternative methods, the VOICE is considered to be ready for operational use. In order to facilitate operational implementation, a contract is planned for FY 1979 in which a computerized vocational counseling system based on the VOICE will be developed. Such a system will allow for rapid computer administration and scoring of the VOICE and will provide immediate feedback to the respondent. The system will be designed to allow the use of alternative scoring methods and adaptive testing strategies and should eliminate much of the bias inherent in paper-and-pencil instrumentation. Whether computer scored or hand-scored by a recruiter, the VOICE should be a valuable counseling tool in the assignment of women to both traditional and non-traditional careers.

Technical training of Air Force personnel represents a considerable investment of financial and human resources, some of which are lost through attrition. Although more women, relative to their proportion in the total force, graduate from BMT than do males, the reverse is true for technical training. Research is now in progress (Kantor et al., 1977a, 1977b) to probe the factors other than aptitude, which may be related to performance in technical school. Specifically, an instrument was constructed to tap student attitudes about motivation, class work, instructors, fellow students, physical aspects of the training enviornment, and satisfaction with the training experience. This Technical Training Student Survey (TTSS) was administered to 11,000 male and 1,800 female students attending five Air Force technical training centers. When the results were analyzed, interesting sex differences appeared. Females indicated more concern about academics such as a desire for more off-duty time and desire to spend more time on difficult subjects. This may be related to their higher attrition rates, but may also reflect a desire to perform at a very high level even if additional time might be required. Females were also less satisfied with certain aspects of the physical enviornment such as classroom temperature and dorm sleeping facilities but had a more positive perception of their classmates. Finally although females seemed happier with their military status in terms of being more satisfied with the Air Force and less bothered by military bearing requirements, they did not feel that technical training had been as beneficial an experience as did males. The instrument used in the technical training study was found to have suitable psychometric properties, was sensitive to student attitudes, and was capable of isolating differences in perception between graduates and non-graduates.

Many of the attitudes expressed by the female technical school students might be considered typical of a group entering a new environment, and, with time, some of the sex differences might be attenuated. However, the information derived from the survey responses can be used as a basis for modifications to the training environment in order to ease some of the problems which make the training experience less satisfying to the female contingent.

The Security Specialist area (AFSC 811SO) is a specialty to which women are being assigned on a test basis only, and the Personnel Research Division is monitoring the progress of 199 females given security specialist assignments in March 1977. For a meaningful evaluation of this pilot program, various groups of accessions were specified for comparative purposes. A control group of 199 male enlistees receiving guaranteed assignments to the 811XO career field were selected by matching with the 199 females on variables such as race, aptitude, AFQT category, education, and age. Additional groups entering service during the same period were identified for further comparisons, i.e., male and female enlistees entering the 811X2 Law Enforcement AFSC and all male and female accessions, regardless of career assignment. A group of 4,300 male and 300 female law enforcement trainees were administered a battery of tests, and their careers are now being monitored to determine the predictive utility of the measures which include biographic/demographic, personal interest, and opinion information. In addition, the same battery of tests was administered to 1,800 male and 120 female security police already on the job. Supervisor ratings of job performance were also obtained. Although only limited data are now avaliable, it was found that no differences existed between the job performance ratings of males and females. Also, similar positive and negative perceptions of the job environment were obtained from both sexes. Further analyses should provide pertinent information pertaining to the performance of males and females in stress situations.

Recently, the Personnel Research Division of AFHRL has begun several research projects to identify possible male/female differences and similarities. One area of research involves psychomotor ability which has been found to be correlated with performance in career fields such as mechanics, electronics, and flying. Using newly developed electronic equipment, the psychomotor performance of both enlisted and officer female personnel is being studied on a wide range of tasks. These data will be compared with extant male data and assessed for similarity in terms of performance and predictive utility for later on-the-job success. Present indications suggest that little or no differences exist in the psychomotor abilities of males and females.

In August 1976, the Air Force began training the first class of female pilots. Prior to training, the Personnel Research Division tested these women on a variety of psychomotor and other tasks found to be associated with success in pilot training. In January 1977, similar data were also collected for the second class of female pilots. Values obtained from subjects of both classes compared quite favorably to similar data previously obtained from male subjects. Of particular interest is female performance on the Automated Pilot Aptitude Measurement System (APAMS). This experimental pilot screening procedure consists of a 5-hour standardized work sample. Subjects fly a GAT-1 simulator through a series of maneuvers ranging from simple straight and level flight to complex traffic patterns complete with strong crosswind effects. A mini-computer montiors the subject's performance and produces a final score reflective of his or her error amplitude on critical flight parameters. In terms of both mean scores and variability, no significant differences were found between male and female scores. As training criteria mature, additional analyses will be conducted, and another group of female pilot candidates was tested prior to training in January 1978 with a revised experimental battery.

Two contract research efforts are related to the utilization of women in the Air Force. One contract attempted to investigate the civilian utilization of women in industrial career fields, and data were to have been collected from 25 corporations. However, due to lack of OMB clearance, data were collected from only three companies. Based on what little data are available, it appears that the Air Force is at the vanguard in terms of policies and practices related to female utilization. The second contract is evaluating male/female differences in civilian aviation. The technical objectives of this project are to investigate type of accident and frequency by sex to determine if either sex appears disposed towards a greater percentage of accidents or types of accidents; to identify and assess differences in instructional needs, student performance and potential; and to identify interest in flying. These data will be obtained through several sources including survey administration to FAA certified flight instructors and pilots, and close investigation of accident reports.

Two additional studies of relevance to the effective use of women in the Air Force are planned for FY 1980 and beyond. One study will investigate the impact of attitudes toward opposite-sex work group members on organizational effectiveness and career development. Under contract, attitude assessment instruments will be developed for administration to incumbents, supervisors, and co-workers in work groups involving participation by opposite sex personnel. The purpose of this research is to determine those attitudes related to gender which serve to limit the satisfaction, productivity, and career development of Air Force personnel. The other study will be concerned with attitudes toward, and the effectiveness of, female supervisory personnel and will not be pursued until more women have attained supervisory positions.

REFERENCES

Alley, W.E., Berberich, G.L., & Wilbourn, J.M. *Development of factor-referenced subscales for the Vocational Interest-Career Examination.* AFHRL-TR-76-88, AD-A046 064. Brooks,AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory, June 1977.

Alley, W.E., Wilbourn, J.M., & Berberich, G.L. *Relationships between performance on the Vocational Interest-Career Examination and reported job satisfaction.* AFHRL-TR-76-89, AD-A040 754. Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory, December 1976.

Berger, F.R., & Berger, R.M. *Vocational Interest-Career Examination: Norming and standardization on a nationwide high school sample.* AFHRL-TR-77-69, AD-A047 762. Brooks AFB, TS: Personnel Research Division, Air Force Human Resources Laboratory, September 1977.

Guinn, N., Johnson, A.L., & Kantor, J.E. *Screening for adaptability to military service.* AFHRL-TR-75-30, AD-A014 790. Lackland AFB,TS: Personnel Research Division, Air Force Human Resources Laboratory, May 1975.

Kantor, J.E., Vitola, B.M., & Guinn, N. *Development and validation of the Air Force Technical Training Student Survey.* AFHRL-TR-77-27(1), AD-A042 967. Brooks AFB, TS: Personnel Research Division, Air Force Human Resources Laboratory, June 1977. (a)

Kantor, J.E., Vitoal, B.M., & Guinn, N. *Male/ female attitudes related to performance in Air Force technical training.* Paper presented at the annual convention of the Military Testing Association, San Antonio, TX, 1977. (b)

SATISFACTION AND EFFECTIVENESS IN THE MIXED GENDER SMALL UNIT

Clay E. George and H. Wayne Smith
Texas Tech University

Three experiments are reported. In Experiment 1, 40 mixed gender triads of students performed a simulated air crew task. There were no gender-related performance differences, but members recorded relatively low satisfaction with the female led crews. Experiment 2 was a partial replication with a traditionally female task and members reported feeling less comfortable with female than male leaders. In Experiment 3, 17 female and 17 male led triads were given a task which provided feedback to members concerning the leader-confederate's effectiveness. Both female and male members expressed increasing satisfaction as performance increased.

Figure 1 represents a model of small unit effectiveness (after George, 1977). One implication of the model is that, when unit members are interdependent and when they receive feedback verifying that interdependence both unit performance and symptomatic variables tend to approach optimal levels. Experimental data upon which the model is based showed that minority (black) members of a group gained in status relative to majority members when and only when conditions of interdependency with feedback obtained (George, 1966; 1977). If female led groups carrying out traditionally male missions experience relatively low levels of symptomatic variables (cohesion, satisfaction, etc.), then training or administrative inputs designed to produce member awareness of leader-member interdependency and effectiveness should produce higher levels of symptomatic variables.

Experiment 1 in the present series randomly assigned 120 university students to one of four conditions. Each condition included ten triads composed of: 1) male leader and two male followers; 2) male leader and two female followers: 3) female leader and two male followers; 4) female leader and two female followers. This procedure resulted in a 2 x 2 x 3 (trial blocks) mixed factorial design. Group members were led to believe that the leader was selected on the basis of objective criteria. Subjects served as pilot (leader) or flight engineers (members) on a simulated air crew problem. Crew members affected an "override on a malfunctioning hydraulic system" on each of six trials by solving a puzzle (Bavelas, 1950). Member communication was allowed only between a member and the leader.

There were no significant differences between gender of leader or gender of followers in performances. There was a main effect showing improved performance over trials; $F(2,72) = 19.67$, $p < .001$. Male led groups, regardless of gender of followers, exhibited $F(1,116) = 4.56$, $p < .05$. *Since the model calls for satisfaction to be determined by performance level, it was suspected that members of female* led crews may have not received sufficient feedback about group performance to realize that their leaders were effective in producing good solution times. Given this situation, the followers may have reasoned that the female leader had the key role in a traditionally male task and was probably not helping the group as much as a male leader would have.

Experiment 2 was a replication of the above except that the crews worked on traditionally female tasks. In this case, the leader took the role of editor of a woman's magazine by helping the members produce good titles for short stories of female interest. Each group produced titles for six stories. There were no gender effects on satisfaction with the group. Male led crews did, however, indicate that they felt more comfortable doing the task than did the female led crews; $F(1,116) = 7.63$, $p < .01$.

The results of these experiments (Smith and George, 1978) indicate that at least a subtle "male leader" stereotype exists which affects symptomatic variables when female leaders are appointed over mixed gender groups. The effect is more general in male stereotyped tasks, but may surface even in female stereotyped tasks. In any kind of task, low levels of satisfaction or comfort with fellow workers may retard the development of intragroup coordination and in the long term, performance, according to the model. These considerations led to an experiment with mixed gender groups doing a task designed to provide feedback to group members concerning their leader's effectiveness. In this situation, the model predicts no gender effects but positive correlations between group performance and satisfaction.

Method

The initial part of the experiment was based on 16 female led triads. Each group had one male and one female member in addition to the leader. Since member gender had no effect on any dependent variable, the experiment was expanded to include 17 male led and 17 female led triads with no attempt to control member gender. Results were thus based on 102 students serving in 34 triads.

The task selected was adapted from Hollander (1960). Each student was given a matrix of five columns and four rows. Cell entries were numbers ranging from -10 to +10. The task was to select a row which would provide a positive payoff when the experimenter announced a column. One could, presumably, work out a rule which would optimize one's payoffs over the nine trials. Each subject decided on a response individually, and then the triad agreed on a group response. Members heard the column number which they used to look up and record the group's payoff trial by trial. At the end of trials 3, 6, and 9, members rated their satisfaction with the group.

Group leaders were appointed by the experimenter as being persons known to have the capacity to help the group perform will. Unknown to the members, leaders were also experimental confederates. They were instructed to inform their members during the first three trials that they knew how to work out the rules involved and would be able to lead the group to obtain high payoffs in the long run. During later trials they were to aver that they had beaten the system and if the group followed their lead, success would be assured. The leader-confederate had been instructed to try to sell the group on responses which would produce payoffs of -4, -1, 0, +2, +6, +7, +8, and +10 respectively, for each of the nine trials. These procedures provided for each three-trial block, measures of group payoff, successful influence attempts with a possible range of 0 to 3 and satisfaction ratings on an 11 point scale.

Results

Leader gender did not affect any of the three dependent measures as a main effect or in interaction, with trial blocks. Payoff increased significantly over trial blocks, as expected; $F(2,66) = 229.51$, $p < .001$. This test was, of course, primarily a manipulation check. Means and SDs were -.085 (5.21), 9.79 (3.92) and 23.76 (5.59) respectively over the three trial blocks. The hypothesis that satisfaction would

increase over time was also supported; F (2,66) = 105.92, p < .001. From Block I to Block III, the means and SDs were 5.56 (2.36), 7.26 (2.05), and 8.58 (1.60), respectively. Leader influence, measured by the number of trials in which the group consensus was the leader's (preinstructed) choice was fairly consistent over trial blocks with some slippage on the last block: F (2,66) = 6.56, p < .05. Means and SDs 2.32 (.72), 2.44 (1.09), and 1.88 (.96).

The most important deduction from the model, however, was that there would be a positive correlation between the (manipulated) payoff scores and the satisfaction ratings. The values for the trial blocks were (32) = .52, p < .01; .15, n.s. and .40, p < .05. In two of the three tests, the prediction was supported that satisfaction increases would follow the manipulated increases in payoff regardless of leader gender.

Discussion

The data suggest that problems sometimes may arise when female leaders are first assigned to leadership positions in mixed gender groups, especially when unit missions involve traditionally male tasks. It is probable, however, that such problems can be abated by providing feedback to members about group effectiveness generally and the adequacy of the leader specifically. It has been pointed out elsewhere (George, 1966; 1977) that unit training can be designed so as to induce symptomatic variables (cohesion, satisfaction with the group) and to optimize unit effectiveness. The present data suggest that higher level leadership can prevent problems in mixed gender units by providing feedback to the members concerning the leader's capacity to generate positive payoffs for the group.

REFERENCES

Bavelas, A. Communication patterns in task-oriented groups. *Journal of the Acoustical Society of America*, 1950, *22*, 725-730.

George, C.E. Training for coordination in rifle squads. *Proceedings: U.S. Army Human Factors Research & Development Conference,*, 1966, *12*, 80-86.

George, C.E. Testing for coordination in small units. *Proceedings: Military Testing Association, 1977*, *19*, 487-497 (microfiche).

Hollander, E.P Competence and conformity in the acceptance of influence. *Journal of Abnormal and Social Psychology*, 1960, *61*, 365-369.

Smith, H.W., and George, C.E. Performance, satisfaction and gneder in appointed leader situations. Unpublished paper. Texas Tech University, 1978.

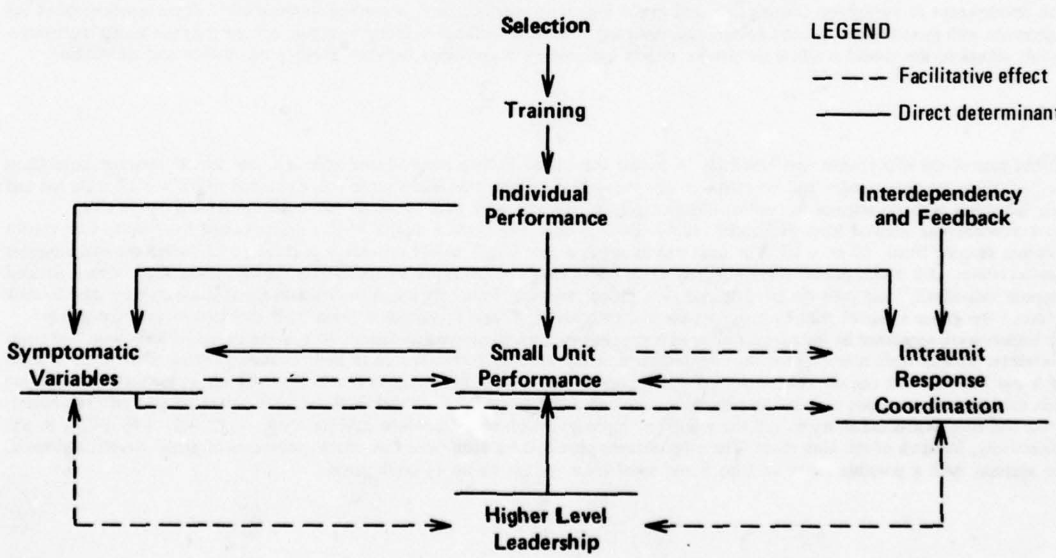


FIGURE 1. A MODEL OF SMALL UNIT FUNCTIONING.

MILITARY PROFESSIONALS /MILITARY LEADERS

Albert R. Brownfield, III
United States Military Academy

Introduction

Two of the most overused words in Army jargon are "professional" and "leader". When listening to soldiers who have passed their initial obligated tours, one could easily get the impression that virtually all of them are both. "Professional" often seems to be connected with time in service or an intention to remain in uniform long enough to qualify for retirement. Whether or not one is a "leader", under the common usage, appears to be a function of rank. This paper begins with the supposition that the manner in which these two terms are misused confuses their meanings and serves no practical end other then the enhancement of the self-esteem that soldier "professionals" and soldier "leaders" feel for themselves. Used in this fashion they do not encourage any predictions about how such soldiers will behave.

This need not be the case. This paper will attempt to define the two concepts, explore the extent to which professionalism and leadership may vary in their occurance among army *officers*, and attempt to predict the degree to which officers who may be effective leaders exhibit some typical professional characteristics.

The Professional

Professions have been variously defined in the literature in terms of their essential traits (Greenwood, 1957; Gross, 1958; Slocum, 1966; Gross, 1966; Perrucci and Gerstle, 1969; Pavalko, 1971) and the negotiation process by which they become professionalized (Friedson, 1970; Roth 1974). A key issue in this negotiation (professionalization) process has been the degree of autonomy permitted the professional group. These two approaches have been linked by professionals justifying their demands for autonomy by citing evidence of the many important "professional" traits in their own activities (Roth, 1974). Such claims to professional traits may be nothing more than public relations devices. This paper will examine a belief that claims to professional status, particularly the claim for autonomy, may paradoxically arise from performance problems. Huntington (1957) and Janowitz (1960, 1970) suggested traits specific for military officers that will be used in this study. To their traits of expertise and responsibility will be added the previously mentioned trait/goal of autonomy in attempting to define military officer professionals.

Leadership: A Critical Performance

Leadership has been examined from the trait perspective (Stogdill, 1948), situational perspective (Leavitt, 1951; Bavelas, 1960; Lawrence and Lorsch, 1969; Hollander and Julian, 1969; Taylor and Bowers, 1970; Jacobs, 1971), and the interactionist perspective summarized by Bons (1976). Of the three the interactionist approach has the most intuitive appeal. At this point it should be specified that, following Henderson and Campbell (1972), leadership is defined as the process of influencing human behavior so as to accomplish the goals of the organizationally appointed leader. Bons' work operationalized this definition by arguing that effective military leaders were those who believed they should serve as a communications link between superiors, peers, and subordinates; who believed in accepting as input to their decision-making process the opinions of all the other relevant actors in the environment; and who believed in recognizing their subordinates as individuals.

In order to summarize the argument on a link between professionalism and leadership the following hypothesis is offered:

H_1 The more an individual believes he should be a communications link between his superiors, subordinates and peers, considers their opinions in his decision making, and believes that subordinates are individuals to be treated as such, the less he will claim that society should give him autonomy in his work.

Other hypotheses attempt to control for various intervening variables that might distort the basic relationship. They address military education:

H_2 If the affects of military education are controlled, the association between the degree to which an individual ascribes to beliefs appropriate to good leadership and the degree to which he claims to base his work on a specialized body of knowledge will be negative.

. . . civilian education:

H_3 If the affects of civilian education are controlled, the association between the degree to which an individual ascribes to beliefs appropriate to good leadership and the degree to which he claims to base his work on a specialized body of knowledge will be negative.

H_4 If the affects of civilian education are controlled, the association between the degree to which an individual ascribes to beliefs appropriate to good leadership and the degree to which he claims to serve the needs of society will be negative.

. . . and the type of assignments an officer has previously completed:

H_5 If the affects of the number of years an individual has been assigned to troop or troop staff duty is controlled, the association between the degree to which he ascribes to beliefs appropriate to good leadership and the degree to which he claims to base his performance on expertise or specialized knowledge will be negative.

One should note that regardless of an officer's assignment pattern or military or civilian education, it should still be expected that individuals with lesser amounts of desirable leadership beliefs would feel strongly that they and the military in general should be given more autonomy by society. Good leaders will rarely need to claim autonomy. Poor leaders may use it to camouflage problems with their performance.

Findings

The hypotheses were tested with results from a questionnaire prepared and distributed to 150 commissioned Army officers stationed at an active army installation. This post provided a particularly heterogeneous sample population in terms of both rank and military specialty. While it had no "troop units" in the conventional sense, it did have two headquarters organizations whose peculiar missions required them to be staffed by officers drawn from all branches of the Army.

The survey questionnaire first asked the respondents to record such basic background information as military rank, component, branch, source of commission, years in service, and levels of military and civilian education. Section two contained twenty-four statements designed to assess respondents' beliefs about leadership and the extent to which they thought army officers exhibited the professional characteristics. Lickert scales permitting a range of responses from 0 to 7 were used with each statement.

Generally speaking, most respondents did not appear to recognize the existance of some special body of knowledge peculiar to their occupation as Army officers. As a group, the respondents did not place themselves at the "highly" professional end of the professional continuum. As a group they did subscribe to the leader beliefs more than to the characteristics of the professional.

A zero order correlation test revealed that the professional characteristics were not highly correlated with one another. Consequently it would be improper to sum all of these characteristics into a single variable and label it "professional".

With the exception of the relationship between consideration of opinions and serving as a communications link the leadership dimensions were also not highly intercorrelated. This suggests two things. First, the three separate beliefs cannot be combined into a single variable. Second, while many officers ascribe to one or two beliefs included in the Bons model, very vew officers ascribe to all three. If one can assume a link between appropriate leadership beliefs and appropriate leadership behaviors, then these results are intuitively sound. Extremely good leaders (individuals with high scores on all three dimensions) and extremely poor leaders (individuals with low scores on all three dimensions) should be quite rare.

Other zero order correlations indicated as predicted predominantly negative relationships between all leadership beliefs and the demand for autonomy. The relationships with the other vo professional characteristics, however, were more mixed than was anticipated. The links between considering others' opinions and the expertise and service dimensions were almost totally independent as was the link between considering subordinates as individuals, and the characteristic of service to country. Considering subordinates as individuals, however, was related negatively to claims of expertise as predicted. The relationships between serving as a communications link and the expertise and service characteristics were quite strong and are in the direction predicted.

A partial correlation test confirmed the first hypothesis with respect to the autonomy claim. The partials further indicated that those officers most likely to demand autonomy in their work were those who did not consider thier subordinates as individuals.

Predictions of beliefs about service and expertise were surprisingly similar. In each case when the positive affects of the communications dimension were removed, the negative affects of considering opinions and considering subordinates as individuals were more clearly visible. These results suggest that with the exception of the communications dimension, the other leadership beliefs tend to correlate negatively with the professional characteristics. The relationships above held no matter which other leadership beliefs were controlled.

It was hypothesized in H_2 and H_3 that the affects of military and civilian education might be surpressing negative links between leadership beliefs and professional claims. Since these relationships were negative to begin with, controls for education should lead to even larger negative relationships. The results indicated most clearly that this was *not* the case. Hence, both the second and third hypotheses were rejected. There was, however, a significant negative relationship between civilian education and a demand for autonomy.

The fourth hypothesis speculated that civilian education might be surpressing negative links between leadership beliefs and professional characteristics. Again, this was not the case. The negative prediction from consideration of opinions increased when education was controlled. The best predictor of the service to a country belief was the extent to which an officer believed he should serve as a communications link.

The fifth hypothesis speculated that the number of years during which an officer was assigned troop or troop staff duty would act as a positive influence on the extent to which he perceived his performance as based on knowledge and expertise specific to his activities. Hence, if the number of years is controlled negative relationships should emerge between leadership beliefs and claims of expertise. The results indicated that controlling the variable had no effect on expertise. In conjunction with the other findings one must conclude that neither type of education nor the number of years of troop service has any effect on perceptions of professional expertise.

When the relationships between demand for autonomy, the leadership beliefs and the number of years assigned to troop duty were tested together, the partials of the leadership beliefs were again unaffected. These relationships remained negative regardless of the number of years of troop duty.

Conclusions

This project began by suggesting that the words "professional" and "leader" were commonly misused by people in the military. The Army officer sample here studied may very well be members of a unique occupational group, but their uniqueness does not cause them to perceive themselves as "professional" if by this is meant that to the same degree they all claim autonomy, all believe they have expertise based on specialized knowledge, and all claim a service of country orientation. Gradford and Brown (1972) argued that the key to military professionalism was the idea of service to country. The relatively high score (5.3) the sample produced on this single characteristic tends to support this much more limited conception.

Similarly, it cannot be concluded that individuals in the sample would uniformly accept all the beliefs implicit in the Bons model of leadership. The low correlations among the various leader beliefs argue that what is recognized as aspects of effective leadership varies from person to person. Yet the average scores on the separate leader beliefs are high. This may reflect the fact that members of this sample have learned that they must remain sensitive to these central issues of effective leadership.

Several times the findings of a consistently negative relationship between a high demand for autonomy from civilian input in military affairs and acceptance of the various leader beliefs were observed. This is interesting and suggestive, but one must use caution in its interpretation. If one accepts the assumption that rejection of the leader beliefs adversely effects performance as a leader, then demands for professional autonomy may indicate performance problems in leadership itself as well as problems in other critical performance areas.

The fact that military education has no measurable impact on any dimension analyzed and does not even moderate the relationship between the various leader beliefs and a perception of some specialized knowledge is indeed surprising. The officers in this sample were not "under-educated". Forty-two percent of them had graduated from the Command and General Staff College. However, whatever else this military education accomplished, it *did not* inculcate a belief that the occupation of Army officers is based upon a specific and unique body of knowledge. Again, the claim to possess specialized knowledge may, like the demand for autonomy, only be made to draw attention away from performance problems.

With the exception of the belief in serving as a communications link, the perception that one possesses a specialized body of knowledge does not relate significantly with any of the other effective leader beliefs. This may indicate that the strong acceptance of the leader beliefs that have been observed are more a product of the on-the-job experiences of the sample than they are of their military education.

An underlying goal of any such study is to demonstrate a means by which certain behaviors might be predicted. With the few exceptions noted above, the data do not encourage such predictions. While not permitting definite conclusions, some interesting relationships have emerged. In the sample the level of civilian education did vary negatively with a demand for autonomy while at the same time varying positively with the notion of being motivated by service to country. Years spent in troop duty assignments also varied negatively with a demand for autonomy. These relationships are sufficiently intriguing to warrant the attention of both Army officers and other analysts.

COMMITMENT PERCEPTIONS OF USAF OFFICERS

V.P. Luchsinger
College of Business Adm.
Texas Tech

Concern for the vitality of organizations is shared today by academicians, behavioral scientists, and members and managers of organizations. Special interest is attached to the roles and behavior of key perssonel in organizations, such as managers, specialists, and staff members. Most large organizations, including the armed forces, are able to attract talented and highly motivated persons. A difficulty occurs in motivating and holding those same personnel over the longer term.

Buchanan (1974) has provided some evidence that commitment is a key variable in retaining personnel. He has further shown that business organizations are more successful in getting and holding talented personnel than are government agencies. Consistently dramatic findings indicate that business executives exhibit more positive attitudes and commitment toward their organizations than comparable government executives.

Traditionally, a crucial role in acquisition and retention of managerial personnel has rested with job satisfaction. Job satisfaction has been a focal point in the role of managers as key decision makers in guiding and executing the missions of organizations. The review of manager job satisfaction has shown findings such as the Paine et al. (1966) study reflecting considerably less government middle manager need satisfaction than similar managers from private industry. More recently, Rhinehart et al. (1969) found dissatisfaction greater among government managers, but that satisfaction tended to increase with both business and government managers with increasing hierarchial levels.

Buchanan (1977) has very recently reported significant differences between government and industrial managers regarding work involvement, job satisfaction, and commitment. Commitment is attracting increased attention in the research literature, although commitment is still a relatively new construct to be considered with regard to job performance and managerial behavior. As managers and organizational scientists and analysts examine new ways of improving employee retention and performance, commitment appears to be shedding some light on methods of upgrading organizational effectiveness.

Commitment is generally viewed to consist of components of: (1) a strong belief in and acceptance of the organizations goals and values (identification); (2) a willingness to exert considerable effort on behalf of the organization (involvement); and (3) a strong desire to maintain membership in the organization (loyalty) (Porter et al. 1974).

Several personal variables have been found to relate to organizational commitment. As attributes of the individual, age (Hrebeniak, 1974) role tension (Hrebeniak & Alluto, 1972), age and central life interest (Dubin, Champoux, & Porter, 1975) have been found to bear on commitment.

Similarly, characteristics of the job have been found to have impact on commitment and job satisfaction, a related affective characteristic. Job challenge (Buchanan, 1974), opportunities for social interaction (Sheldon, 1971), and feedback on the job (Porter & Steers, 1973) have been tied to commitment.

The quantity and quality of relationships to the work organization have also been held to contribute to commitment. As a socializing force Buchanan (1974) claimed psychological attachments to the organization were seen as powerful contributors to commitment. Other social influences were organizational trust and dependability (Hrebeniak, 1974), perceptions of personal investment and personal importance to an organization (Buchanan, 1974), and realization or expectation of rewards in the organization (Grusky, 1966).

The increasing importance of commitment to organizations is attributed to the ability of commitment to better predict turnover than job satisfaction (Porter et al., 1974), and as a contributor to effectiveness of organizations (Steers, 1975). Steers (1977) has recently found that commitment is related to categories of personal characteristics, job characteristics, and work experiences, while also related to the stability of the work force.

This study represented a pilot study to examine commitment among a group of Air Force officers engaged in graduate studies in business administration. The Porter (1974) framework was followed in measuring the extent of commitment expressed on a Likert 7-point scale instrument. A modification for this study shredded out commitment to: commitment to the organization (USAF), commitment to the work group (primary) and commitment to leadership. The purpose was to examine differences in perceived commitment between the organization, work, group, and leaderhsip as related to job satisfaction, rank (with tenure implied), role clarity, and propensity to leave.

Table 1 shows the relative strengths of perceived commitment to the various objects of commitment. In general, the higher ranks showed more commitment to organization, work group, and leadership. The combined groups were most commited to work groups, then to the

TABLE 1
PERCEIVED COMMITMENT TO ORGANIZATION, WORK GROUP, AND LEADERSHIP AS A FUNCTION OF RANK

	Organization	Work Group	Leadership
Captain and below N = 41	3.1	2.2	5.3
Major and above N = 19	2.0	1.3	3.1

(1 = highest commitment, 7 = lowest commitment)

organization, and finally to leadership. As with other findings, these results seem to suggest that participants with more tenure and age perceive more commitment since they are still in the system, and possess an investment in the system.

To junior members, they represent the system. Junior officers were most critical of leadership, which represented an impetus for leaving the system. The junior officers mean job satisfaction was 3.3 compared to 4.2 (on a 1 to 7 ascending scale) for the senior officers. Junior officers were more inclined to leave the organization, with the organization being the chief irritant, then leadership being impetus to leave. By contrast the senior officers were bothered by both leadership and the organization as impetus to leave.

TABLE 2
PERCEIVED CONTIRUBITORS TO JOB SATISFACTION AS A FUNCTION OF RANK

	Organization	Work Group	Leadership
Captain and below	4.0	2.6	3.5
Major and above	2.4	2.1	3.5

Table 2 shows that affiliation with the work group was the strongest contribution to job satisfaction. Leadership contributed least for the senior officers, and the organization least for the junior officers. As a preliminary study, this work may point to some implications for advanced examination of commitment as an organization phenomenon.

For instance, commitment can become an organizational and personal attribute above and beyond job satisfaction. In effect, some persons may become and remain committed to work groups, leadership, and or organizations despite job dissatisfaction. In other cases, job satisfaction could prevail with or without commitment.

Commitment is worthy of more study. Salancik (1977) has pointed out that the goal is to get the individual committed to behaviors that are right for him and right for the organization. Committed persons stay on the job, they require less supervision, and are able to enjoy the all too rare congruence of individual and organizational goals.

Buchanan (1975) has concluded that organizations can foster commitment by: designing jobs tailored to skills of incumbents and related to larger goals, communicating expectations for commitment, and by designing career development structures in which incumbents receive feedback about progress.

This study shows on a preliminary basis that building and maintaining cohesive work groups could be a key to commitment, when they are the objects of the strongest commitment of individuals. This rationale would be further support for enhancement of organizational development (O.D.) efforts.

BIBLIOGRAPHY

Buchanan, B. Government managers, business executives, and organizational commitment. *Public Administration Review*. July-August 1974, 339-347.

Buchanan, B. Building organizational commitment: the socialization of managers in work organizations. *Administrative Science Quarterly*, 19: 533-546.

Buchanan, B. To walk an extra mile: the whats, whens, and whys of organizational commitment. *Organizational Dynamics*, Spring, 1975, 67-80.

Dubin, R, Champoux, J, & Porter, L. Central life interests and organizational commitment of blue-collar and clerical workers. *Administrative Science Quarterly*, 20: 411-421.

Grusky, O. Career mobility and organizational commitment. *Administrative Science Quarterly*, 10: 488-503.

Hrebeniak, L. Effects of job level and participation on employee attitudes and perceptions of influence. *Academy of Management Journal*, 17: 649-662.

Hrebeniak, L. & Alutto, J. Personal and role-related factors in the development of organizational commitment. *Administrative Science Quarterly*, 17: 555-572.

Pearce, J. An assessments of supervisor's loyalty. *MSU Business Topics*, Summer 1977: 50-56.

Porter, L., Steers, R., Mowday, R., & Boulian, P. Organizational commitment, job satisfaction, and turnover among psychiatric technicians. *Journal of Applied Psychology*. 59: 603-639.

Porter, L. Crampon, W., & Smith, F. Organizational Commitment and Managerial Turnover. *Organizational Behavior and Human Performance*, 1976.: 87-98.

Salancik, G., Commitment is too easy. *Organization Dynamics*, Summer 1977: 62-80.

Sheldon, M., Investments and involvements as mechanicsm producing commitment to the organization. *Administrative Science Quarterly*, 1971. 16: 142-150.

Steers, R., Antecedents and Outcomes of Organizational Commitment. *Administrative Science Quarterly, 1977.* 22: 46-56.

Paine, F., Carroll, S., & Leete, B., Need satisfaction of managerial level personnel in a government agency. Journal of Applied Psychology, 1966.

Rhinehart, R., Barrell, A., DeWolfe, J., & Spaner, F., Comparative Study of need satisfaction in governmental and business hierarchies. *Journal of Applied Psychology*, 1969.

ANTECEDENTS OF ORIENTATIONS TOWARD MILITARY UNIONS AMONG COMBAT TROOPS

David R. Segal
Robert C. Kramer
University of Maryland

One of the consequences of the transformation of America's armed services from a conscription-based force to an all-volunteer force has been an increasing similarity between military service and employment in the civilian labor force (Moskos, 1977; Segal, 1977). Following the recommendation of the President's Commission on an All-Volunteer Force (1970), the armed services have, since the early 1970s, competed with civilian employers for quality personnel, using many of the same recruiting incentives as these employers. A natural outgrowth of this strategy has been an hypothesized transformation of military service from a calling to a job (Moskos, 1977).

As the personnel costs of the all-volunteer force have increased, however, the traditional benefits of military service, such as medical and dental care, education, and commissaries, have increasingly come under attack, leading military personnel to perceive that their benefits were being eroded. This perception, coupled with the increased similarity between military service and civilian employment, suggested that armed forces personnel might follow the precedent established by non-uniformed government personnel, and seek labor union representation in dealing with their grievances (e.g., Krendel and Samoff, 1977). In September, 1976, the American Federation of Government Employees (AFGE) voted to amend the AFGE constitution to make armed forces personnel eligible for membership. Although the AFGE decided subsequently, on the basis of a poll of its locals, not to attempt to organize military personnel, the Congress has been moving foward with legislation to preempt future attempts to unionize the armed forces.

The changing nature of military service as a consequence of the conversion to an all-volunteer force, the precedents set by the progressive unionization of civilians employed in the public sector, the experiences of many of our European allies with unionized armed forces (see Krendell and Samoff, 1977), and the willingness of organizations such as AFGE to consider seeking to represent military personnel, all lend credence to the possibility that the United States armed forces will at some time in the near future, have to deal with active attempts at unionization (see Taylor, Arango, and Lockwood, 1977). The issue we seek to address here is the orientation of military personnel themselves toward unionization, and the factors that affect the orientation.

Military Personnel View Unionization

No body of data representative of all uniformed personnel across the services has been brought to bear on the issue of unionization. Several studies of subpopulations have been undertaken however, and have produced sufficiently consistent findings that we have considerable confidence in the generality of the patterns observed. In a study of a representative sample of Air Force active duty personnel in 1976, Manley, McNichols and Young (1977) estimated that about a third of their sample would be willing to join a union, that there was widespread perception of benefit erosion, but that this was offset by a feeling that unionization would negatively affect the internal operation of the service. Officers saw this impact as more serious than did enlisted personnel.

In another 1976 survey, McCollum and Robinson (1977) similarly found that among "mid-career" non-commissioned officers (E-5 through E-7) and captains at Fort Bragg, North Carolina, about one-third of the sample was willing to join a union. They also found a pattern of concern about benefits and of the impact of unionization on the service similar to the pattern observed in the Air Force. Differences between the officers and the non-commissioned officers in the sample also reflected the pattern observed in the Air Force.

Carlton and Enderlein (1977), using a modified version of the survey instrument used at Fort Bragg, surveyed 221 army reservists. A larger proportion of these personnel indicated a positive propensity toward union membership. This finding is not surprising given the assumptions that unionization is one reflection of the "civilianization" of the military, and that the reservist is one prototype of the "citizen-soldier," whose attitude would be expected to be different from the military careerist in the all-volunteer force. Carlton and Enderlein do not provide sufficient data to determine whether the general patterns observed in the Air Force and Fort Bragg surveys appeared among the reservists as well (for an excellent overview of these surveys, see Pillsbury, 1977).

The Fort Benning Survey

In 1977, a sample of Army personnel at Fort Benning, Georgia, stratified by rank, was surveyed, using a 63 item questionnaire based upon the survey instrument used in the Air Force survey, with minor adaptations for use in the Army. The sample was then weighted to approximate the rank distribution of the total Army. The weighted n was 695. The weighting does not make the sample representative of the Army. Our respondents were all at one post, which contains a major training facility (the Infantry School), and which thus does not reflect a normal garrison situation. Our respondents were all in combat specialties, and they were disproportionately oriented toward military careers (Segal and Kramer, 1977). The data from this survey, however, can be used to provide initial estimates of the magnitude of impact of factors associated with pro-union or anti-union sentiment. The proportion of respondents in this survey who indicated that they would join a union was very close to the one-third figure found in other surveys.

Measuring Orientation Toward Unions

Previous research has focused on the willingness of respondents to join a union. Both because of a high incidence of "dont' know" responses (generally more than 30%), and because of the low reliability of single item measures, we constructed an index of orientation toward unions based upon five items in the survey. We cumulated the scores of our respondents on 5-point Likert-scale items regarding (a) the need for a military union; (b) whether military personnel in non-critical jobs should be allowed to strike; (c) whether the respondent would join a military union; (d) whether military unionization is inevitable; and (e) whether the Army should do everything it can to avoid unionization. All five items had loadings of at least .60 on the first factor in a principal components analysis. The index ranged from 5 to 25, with a mean of 12.9 and a standard deviation of 5.17. Individuals scoring at the high end of the scale were those who indicated

that there was a need for a military union; that they would join one; that military personnel should be allowed to strike; that untionization is inevitable; and that the Army should not try to avoid unionization.

Correlates of Union Attitudes

Previous research has shown that background characteristics such as level of education, prior experience with unions, military rank, number of years of service and level of information about unions are related to views of military unionization. Orientation toward a military career also has such a relationship. We sought to supplement these variables with soldier's attitudes toward a number of issues that might be related to unionization. Four attitude dimensions were suggested by factor analysis of the remaining attitude measures.

The first of the dimensions was concerned with the impact of unionization on the Army. Nine items loaded at .5 or higher on this factor, and were cumulated to form this index. Six of the items were Likert scales, and dealt with (a) whether membership in a military union would result in better working conditions; (b) whether a union would have a negative effect on discipline in the Army; (c) whether a union should have a say in operational matters; (d) whether union representation would insure that military personnel would be treated with dignity; (e) whether a union could negotiate disputes between military personnel and the Army; and (f) whether there is a need for such third party representation. The remaining three items dealt with the implications of unions for (g) relations between superiors and subordinates; (h) military professionalism; and (i) the mission effectiveness of the Army.

The second index suggested by the factor analysis was a measure of job satisfaction. The Air Force survey instrument had included the Hoppock (1935) job satisfaction measure: a four item scale based upon (a) how much of the time the respondent feels satisfied with his job; (b) how well the respondent likes his job; (c) how he feels he compares to other people in this regard; (d) how he would feel about changing his job. Our factor analysis combined these items with two other job satisfaction items included in the survey: (e) the degree to which the respondent saw his present job as boring or challenging, and (f) whether his present job was preparing him to assume positions of greater responsibility. All six of these items had factor loadings of .5 or higher, and were included in our index of job satisfaction.

We had anticipated that job satisfaction would appear as a factor in our analysis, and indeed that it would be related to orientations to military unionization. We had not anticipated the third factor in our analysis: evaluations of the Army's promotion system. This factor was defined by two questions: the respondent's evaluation of whether the military promotion system is fair; and the respondent's evaluation of whether the military promotion system is effective. Recalling that much of the early analysis of soldier attitudes by Stouffer and his colleagues (1949) was concerned with perceptions of promotion opportunities, we retained this factor in our analysis.

We had anticipated that the factor analysis would yield a factor reflecting perceptions of whether soldiers felt that they would benefit from unionization. This dimension did emerge as the fourth factor, with two items loading highly on it: whether military unions could secure higher pay raises, and whether unions obtain more benefits for their members than would be obtained with unions.

Table 1 presents an overview of the four factors extracted from our attitude matrix after the attitudes comprising our dependent variable were deleted. These four factors account for 82.8 percent of the variance in the matrix. Note that we did not use factor scores in building our indexes of these dimensions. We merely cumulated item scores, reversing item scoring where necessary so that all items in a given index scaled in the same direction.

TABLE 1. Factor Analysis of Soldier Attitudes

Factor	Eigenvalue	Percent of Variance
Impact of Union on Army	7.838	45.4
Job Satisfaction	3.080	17.8
Promotion	1.895	11.0
Benefits	1.489	8.6

Analysis

A series of multiple regression analyses were performed to identify the direct and indirect effects of years of service, level of education, prior experience with unions, military rank, career orientation, perceived ability of unions to gain benefits, information about unions, expectations of the impact of unions, job satisfaction, and evaluations of the promotion system, on the orientation toward unions. The results of the basic regression analysis are presented in Table 2.

TABLE 2. Regression Analysis of Factors Influencing Orientations Toward Unions.

Variable	Multiple r	r^2	beta
Career Orientation	.175	.0306	-.093
Impact of Union	.833	.6938	-.790
Information about Unions	.841	.7068	.100
Benefits through Unions	.843	.7113	.079
Education	.848	.7194	-.098

Note that only five antecedent variables were shown to have significant effects at p< using an F-test. Evaluations of the promotion system job satisfaction, prior experience with unions, and length of service, all fall out of the model. Our finding regarding job satisfaction is one more contribution to a growing literature which demonstrates that while we can find antecedents to job satisfaction, it is difficult to identify *consequences* of job satisfaction. While our analysis showed career orientation, military rank, and level of education all to have significant positive effects on job satisfaction (multiple $r^2 = .145$), job satisfaction did not in turn have any effect on orientations toward military unions.

Figure 1 presents the results of our regression analyses in the form of a path diagram. The coefficients shown in this figure are standardized regression coefficients. Although all possible direct and indirect casual paths were analyzed, the path diagram presents only those coefficients that were significant at $p < .05$, using an F-test.

Our model of orientations toward military unions was nonetheless a powerful one, explaining more than 70 percent of the variance in the dependent variable (multiple $r^2 = .719$). Most of the explanatory power was due to the direct effect of expectations regarding the impact of unionization of the Army. Personnel who thought that unionization would not result in better working conditions, would have a negative effect on Army discipline, would negatively effect relations between superiors and subordinates, military professionalism, and mission effectiveness were not favorably disposed toward military unionization. Perceptions of the impact of unionization, in turn, were affected directly by career orientation, by the respondents level of education and military rank. These factors, however, explained a relatively small proportion of the variance in perception of impact (multiple $r^2 = .188$). What is more important in this portion of the analysis is that the majority of the personnel in our sample felt that unionization would disrupt the functioning of the Army, and were therefore not favorable disposed military unionization.

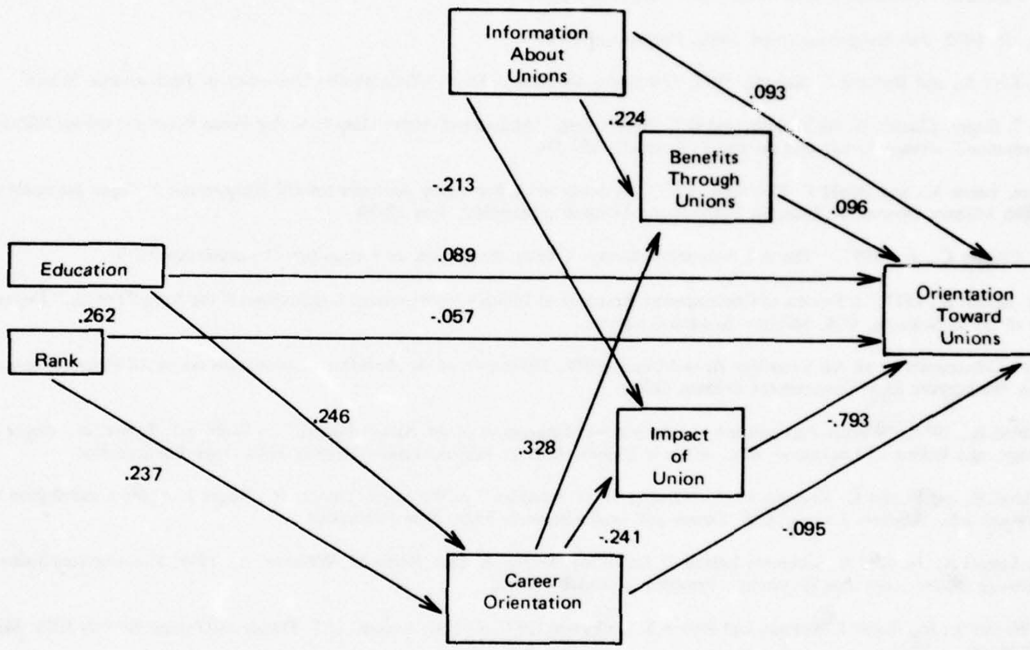


FIGURE 1. Antecedents of Orientation Towards Unions.

Perception of the ability of unions to benefit Army personnel also had a direct effect on orientation toward unions: considerably smaller but statistically significant. The impact of this variable is constrained by truncated range in the perception of benefit erosion. So widespread were preceptions that benefits were being eroded and that third-party representation was needed to counteract this trend that there was relatively little variance in this index.

About 10 percent of the variance in perceptions of the abilities of unions to gain benefits was accounted for by the direct effects of career orientation and level of information, the indirect effects of level of education, and military rank.

Level of information about unions, career orientation, level of education, and military rank all had direct effects on orientations toward unions, as well as exerting indirect effects through intervening variables.

Discussion

Our analysis does not show much enthusiasm for military unionization among active duty combat troops. While pro-union sentiment might be lower in our sample than in the armed forces generally, other data on active duty personnel suggest that the differences are not great. Reserve personnel are another matter. There we would expect the "citizen" component of the citizen-soldier ethic to be related to more positive views toward unionzation.

While support for military unionization is a minority view, it is not to be taken lightly. We regard it not so much as a militant demand for worker democracy in military organization as a sympton of an underlying organizational malaise. Our respondents are telling us that

they feel their service is not valued by the society they are pledged to protect. They see the military being held in low esteem by the American people, and they see their benefits being taken away by the Congress. Most importantly, they see no one standing up to defend them.

Their commitment to duty and service is reflected in the fact that even given their perceptions of the treatment they are receiving, their expectations of the impact of unionization on the Army causes them to oppose such organization. As long as their relative position, as they perceive it, does not deteriorate further, and they continue to believe that unionization will undermine discipline, effectiveness, professionalism, etc., we do not anticipate an increase in support for military unionization. Indeed, if the perceived trend toward erosion of benefits is believed to be reversed, or if they see the Defense Department, the Congress, and the nation manifesting greater appreciation for their contribution, support for unionization of the military will in all likelihood decline. If, on the other hand, they feel that the trend toward benefit erosion is continuing, or that they can organize in ways that will not undermine the functioning of the Army, e.g., through more militant activity on the part of existing service associations, widespread organization of the services is likely. In any case, it is clear from our data that stopgap measures aimed at improving the job satisfaction of military personnel will not deter unionization of the military.

REFERENCES

Carlton, Patrick W., and Maxine A. Enderlein, 1977, "Army Reservists and Military Unions." Paper presented at the 85th Annual Convention, American Psychological Association, San Francisco, August 27.

Hoppock, R, 1935, *Job Satisfaction* (New York: Harper and Row).

Krendel, Ezra S., and Bernard L. Samoff, 1977, *Unionizing the Armed Forces* (Philadelphis: University of Pennsylvania Press).

Manley, T. Roger, Charles W. McNichols, and G.C. Saul Young, "Attitudes of Active Duty U.S. Air Force Personnel toward Military Unionization." *Armed Forces and Socieyt* 3 (Summer): 557-574.

McCollum, James K., and Jerald F. Robinson, 1977, "A Study of Active Army Attitudes toward Unionization." Paper presented at the 39th Military Operations Research Symposium, Annapolis, Maryland, June 28-30.

Moskos, Charles C., Jr., 1977, "The All-Volunteer Military: Calling, Profession, or Occupation." *Parameters* 7:2-9.

Pillsbury, Hobart B., 1977,"A Review of Contemporary Research on Military Views toward Unionization of the Armed Forces." Department of Social Sciences, U.S. Military Academy, mimeo.

President's Commission on an All-Volunteer Armed Force, 1970, *The Report of the President's Commission on an All-Volunteer Armed Force.* Washington: U.S. Government Printing Office.

Segal, David R., 1977, "Worker Participation and the Future Management of the Armed Forces," in William J. Taylor, Jr., Roger J. Aranago, and Robert S. Lockwood, eds., *Military Unions: U.S. Trends and Issues* (Beverly Hills: Sage Publications).

Segal, David R. and Robert C. Kramer, 1977, "Ground Force Attitudes," in William J. Taylor, Jr., Roger J. Arango, and Robert S. Lockwood, eds., *Military Unions: U.S. Trends and Issues* Berverly Hills: Sage Publications).

Stouffer, Samuel A., Edward A. Suchman, Leland C. DeVinney, Shirley A. Star, Robin M. Williams, Jr., 1949, *The American Soldier: Adjustment During Army Life* (Princeton: Princeton University Press).

Taylor, William J., Jr., Roger J. Arango, and Robert S. Lockwood, 1977, *Military Unions: U.S. Trends and Issues* (Beverly Hills: Sage Publications).

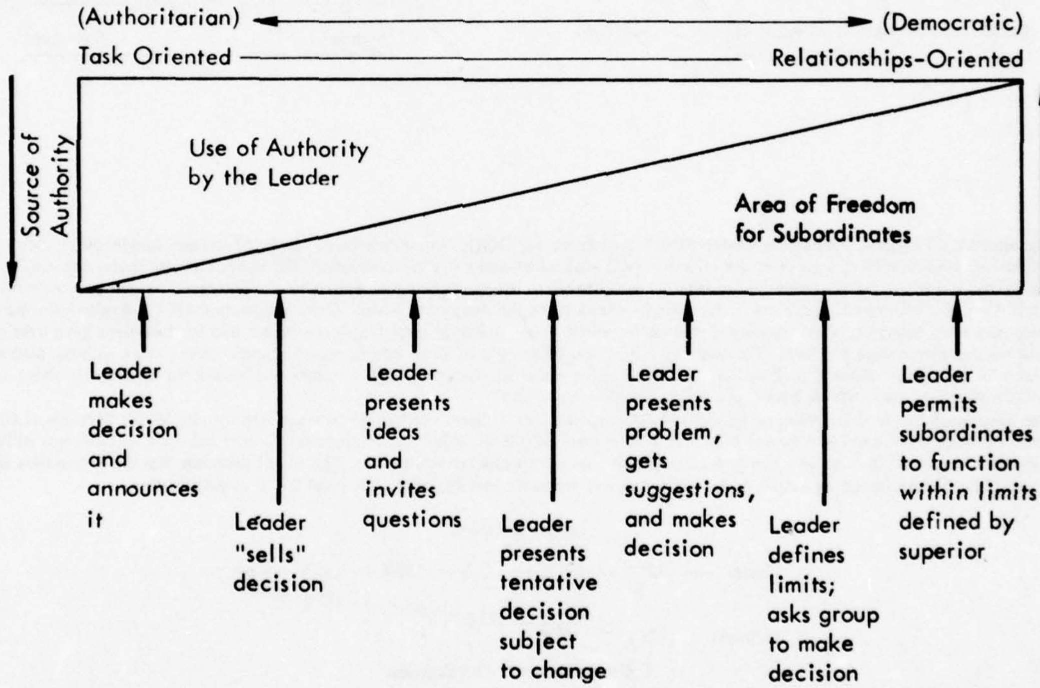
A SITUATIONAL LEADERSHIP MODEL FOR DOD

Loren I Moore
U.S. Navy

This situational leadership model recognizes that a common factor of DOD administration is the requirement for people within organizations to influence the behavior of others, i.e., leadership. Effective situational leadership is a function of the leader, the followers, and the situation: L = (l,f,s,). This interrelation of the elements of leadership does not permit the actual isolation of one element for either research or *in situ* leadership. However, in the education, training, or workshop setting, an element can be emphasized. Individual leadership styles and range may be determined and a Situational Leadership Influence Model (SLIM) developed based upon the relationship between effective Leader Behavior (LB), and the group Influence Development (ID) level of the followers within a situation (Moore 1975).

The literature of leadership research follows the pattern: The Great Man Approach--leaders are born, not made: the Trait Approach--leaders differ in certain characteristics; the Functional Role Approach--leadership exists in a group when various tasks and maintenance functions are performed appropriately; the Continuum or Styles of Leadership Approach--leadership can be differentiated into authoritarian or democratic styles; and, the emerging Situation Approach--the leader can exercise leadership behavior appropriate to the situation. This latter may be the most fruitful approach for those in the Department of Defense.

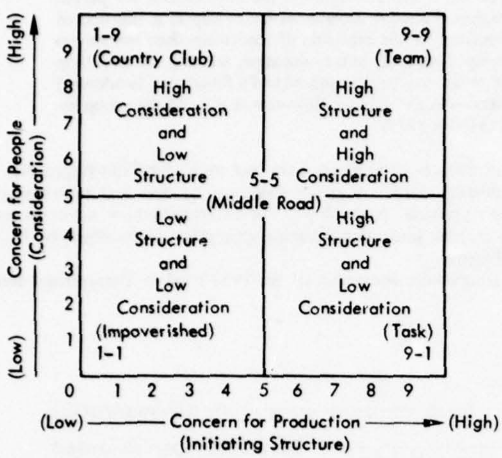
The basic structure of the research on Authoritarian, Democratic, and laissez-faire leadership of the 1930's led to Tannenbaum and Schmidt's, "Choose a Leadership Style" model:

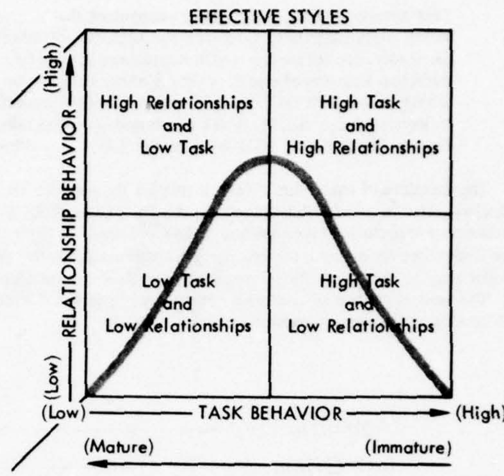


Next came two-dimensional plotting; such as the Ohio State and the Managerial Grid, and The Life Cycle Leadership Theory (Hersey and Blanchard 1972):

The Ohio State and Managerial Grids do not have a situational component and therefore are not appropriate models.

Hersey and Blanchard's life cycle leadership model is an effective, simplistic, situational model. However, it has several drawbacks. The model reads upside down and backwards. Maturity determination, except in the vaguest terms, is not developed by the originators. The very word "maturity" has evaluative connotations. Similarly, the term "life cycle" is presently used within DOD in a completely different contest. Many persons appear to misunderstand the life cycle model and presume that because there is a "lack" of leadership behavior in Quadrant 4, that the task is not being accomplished. The model appears to indicate definitive divisions of leader behavior, and follower maturity. A more realistic approach to situational leadership models might be more appropriately discussed in terms of tendencies and probabilities. Finally, the behavior of the leader is presented in terms of "task" and "relationship" behaviors. The populations upon which these basic distinctions were made were white, middle class males of the 1940 to the 1960 period. Their applicability to the culture of 1977 and beyond may be appropriate; however, there is no specific data to warrant such a conclusion.





The Situational Leadership Influence Model (SLIM) is proposed for DOD. Simple definitions for the SLIM are: Leadership is the process of influencing the acitivties of a group in the efforts toward goal achievement in a given situation. Effective, efficient leadership is a function of the leader, the followers and with the situation L = (l,f,s) towards appropriate goal achievement. The *leader* is that person designated as such. Only the designated leader can exhibit group who are not the designated leader. Only followers exhibit follower influence (ID) development level behavior. The *situation* is that environment or surroundings over which the leader and the followers have little or no control without extraordinary efforts. The situation might best be thought of as an infinite series of concentric circles with the leader and followers in the center. Even though an amount of influence could be passed through a number of situational boundaries, there would be infinite situations over which leader and followers have no control.

The Situational Leadership Influence Model (SLIM) hypothesizes a direct correlation between appropriate leader behavior (LB) and follower behavior manifested in follower influence development (ID) level. While the continuum of leader behavior and follower influence development have been divided into five general segments, such divisions are not distinct. The SLIM proposes that the appropriate leader behavior (LB) is determined or established by the follower influence development (ID) level for a given situation:

Leader Behavior

Leader — LB-1 — LB-2 — LB-3 — LB-4 — LB-5 →

Follower — ID-1 — ID-2 — ID-3 — ID-4 — ID-5 →

Follower Influence Development

Beginning with structuring leader behavior—the appropriate behavior for working with undeveloped follower—the SLIM suggests that the leader behavior should move through LB-1 predominate structuring behavior to LB-2 high structuring and low consideration behavior to LB-3 equally high levels of structuring and consideration behavior to LB-4 high consideration low structuring behavior to LB-5 equally low consideration and structuring behavior as (and if) the followers develop higher levels of influence.

Levels of follower influence development, ID are provided by dividing follower influence continuum into five general areas: *initial, low, average, high average,* and *high*-simply marked ID-1 through ID-5 for a specific task. When leading followers who have just formed as a group (hence *initial* ID-1 stage) a highly structuring leader behavior LB-1 has the highest probability of success; with followers exhibiting a *low* level of influence development ID-2, high structuring and low consideration leader behavior, LB-2, is applicable; for *average* and *high average* development levels, ID-3 and ID-4, equally high structuring and consideration leader behavior LB-3 to high consideration behaviors and somewhat lower levels of structuring leader behavior LB-4 appear to be most appropriate; and low consideration-low structuring leader behavior LB-5 has the highest probability of success working with followers who have a *high* level of influence development beyond ID-5, leader and follower behavior merge to truly collegial behaviors.

Influence development ID level is defined as the degree of achievement motivation, willingness and ability to take responsibility, and task-relevant education and experience of an individual or a group (Hersey and Blanchard, 1972). Additionally, influence development, ID is congruent with changes in behavior from passive to active, from dependent to independent, from the ability to behave in few ways to the ability to behave in many ways, from subordinate to equal or superordinate positions, and from lack of awareness and control to awareness and control over self or the actions of followers (Argyris, 1957).

Each dimension is placed on a continuum assuming, for the follower development purposes, that attributes and aspects of individual personality are applicable to the followers as a group (Likert, 1971). The dimensions are descriptive of basic multidimensional processes through which the development of followers influence may be observed. At any time, any group can have its degree of influence development ID established within these follower influence dimensions.

Follower influence development, ID level, then, is the degree of follower behavior observed in verbal and nonberbal manifestations of these dimensions. Hundreds of persons in a variety of task groups have determined the ID level of their own groups. The analysis of data was conducted by discussions following group tasks and by *post hoc* review of video tapes. Followers in a wide variety of settings, i.e.: operational military units, college faculties, military and civilian curriculum development sessions, civilian and military administrative personnel meetings, degree candidates classes, community action programs, humanistic educational experiences, military civilian executive development sessions, career educational workshops, etc., evaluated their own behavior and established their own group's influence development ID level. Each of the dimensions have been examined with the verbal and nonverbal behaviors reported in detail (Moore 1976).

For example: Achievement. McBer listed an achievement process that could include setting initial goals, anticipating problems and risks, planning and taking moderate-risk actions, obtaining necessary information for task completion, reviewing programs, and revising (McBer 1970). The quality of task achievement appears to increase, or at least remain nearly the same, as followers develop, for the achievement dimension.

The leader observes behaviors, either verbal or non-verbal, in which the followers take individual and group responsibility; seek concrete feedback; attempt creative or innovative solutions; attempt to outperform other groups; attempt to set and meet self-imposed standards; and use power or affiliation to accomplish the assigned task (McClelland, 1961). Only in rare cases can followers be considered developed if they do not achieve the assigned.task.

There is often a conflict between achievement of short-range and long-range goals. As Forrester (1971) points out, short-range goals may be in direct opposition to long-range goals. One good indicator of *average to high average* development ID-3 and ID-4 appears to be when the followers recognize the difference between their short-range and long-range goals.

Probably the most difficult problem when using achievement as a dimension for determining influence level ID is that in many situations there is no single, correct "achievement", (although there may be such correct achievements in education and training situations). Organization development, OD programs try to solve this problem by having everyone share a "common vision" of where the enterprise should be going. Another technique is to create a strategy and then to build a series of guidelines or steps along the way to the goal. The followers can determine how well they are proceeding toward achieving their goal.

High Average and *High*, ID-4 and ID-5 influence development level followers will be effective and efficient. They will expend effort and resources in comparable measure to the goal to be achieved and will recognize the time limitations that are placed upon tasks. It appears that most groups achieve at levels high enough for survival, but well below their potential.

The observable follower influence development ID levels of a newly formed group is virtually always *initial*: ID-1. The most consistently appropriate leader behavior is *initial* LB-1 high structuring behavior. Structuring for task emphasis with newly formed groups would greatly facilitate a solution to many leadership problems. Groups first formed ID-1 or with an observable *low* ID-2 level of follower influence development consistently fail to accomplish tasks if the leader, or another member of the group, does not exercise structuring behavior.

The determination of follower influence development ID level, in even the most general terms, was consistently commented upon by participants in training (and field) situations as an appropriate basis for leader and follower actions. Participants demonstrated changed behavior—attempting to exhibit appropriate leader behavior (LB)—on the basis of observable follower influence development ID levels.

The most common mistake made in the determination of follower influence development ID level is to assume that follower ID level is the sum of the individual's automatically equals a developed group is wrong. The leader then erroneously uses a style of leader behavior considered appropriate for higher levels of development. Because the diagnosis is incorrect, however, the results are less than optimal. The leader should use high structuring behavior LB-1 with newly formed groups, *then* change as rapidly as possible through high structure to low consideration, LB-2,to LB-3, (high structure-high consideration behavior) based on observable follower behavior changing from *initial*, ID-1 to *low*, ID-2 to *average*, ID-3.

LB-1, LB-2 and LB-3 covers the majority of follower influence cases: *initial*, ID-1, low, ID-2, and average, ID-3. Therefore, Leader Behavior, LB-1, LB-2,and LB-3 are the recommended leader's usual styles. With definite observable indications or higher levels of follower influence development, (ID-4, ID-5) the leader could use LB-4 or, rarely, LB-5 leadership styles with sub groups or individuals as appropriate. The least risk to task accomplishment and group morale is to emphasize LB-3 (high structuring, high consideration) leader behavior.

Hierarchical organizations and models, particularly in the military and business, tend to predispose people to think of influence as possessed and exercised only by those in a superior position. The SLIM, with some emphasis on followers, helps indicate that all the elements of leadership possess influence. The determination of follower influence development, ID levels allow the leader (and followers) to utilize more fully the influence of the followers for optimum task accomplishment within DOD

The ability to diagnose organizational behavior, in terms of influence development, ID levels, provides an appropriate basis for personal leader or follower behavior within a DOD system or for recommended interventions as a consultant to a DOD Unit.

Organizations may be viewed as a linking of followers and leaders in various arrays and positions. By systematically establishing the leader and followers in each specific submodule, a determination of follower influence development (ID) can be made. The submodules can then be viewed as a whole and appropriate leader behavior (LB) may be used.

Videotaping follower and leader behavior offers a great promise for systematic observation and analysis of complex influence concepts. A set of procedures to organize follower activity (such as a seminar) allows follower influence to be observed, recorded, and analyzed. To accomplish these ends, the dimensions of follower influence have been developed. These dimensions can then be used to identify, record, and measure the events that take place in the myriad leader follower situations within DOD.

REFERENCES

Argyris, C., *Personality and Organization*. New York: Harper & Row, 1957.

Forrester, J.W., *World Dynamics*. Cambridge, Mass.: Wright Allen Press, 1971.

Hersey, P., & Blanchard, K.H., *Management of Organizational Behavior* . Englewood Cliffs, N.J.: Prentice--Hall, 1972.

Likert, R., *The Nature of Highly Effective Groups*. In D.A. Kolb, I.M. Rubin, J.M. McIntyre (Eds.) *Organizational Psychology: A Book of Readings* (2nd Ed.). Englewood Cliffs, N.J.: Prentice-Hall, 1971.

McBer & Company. *Measuring Motivation*. Cambridge, Mass: Sterling Institute, 1970.

McClelland, D.C., *The Achieveing Society*. Princeton, N.J.: D Van Nostrand, 1961.

Moore, L.I., "Towards the Determination of Follower Maturity: An Operationalization of Life Cycle Leadership." Unpublished doctoral dissertation, University of Mass., 1975.

Moore, L.I., The FMI: Dimensions of Follower Maturity. *Group and Organizational Studies, The International Journal for Group Facilitators*, 1976, 2, 203-222.

THEORIES OF LEADERSHIP

Mickey R. Dansby
AFIT/ CIP, University of Florida

Leadership has been an intriguing subject for the social analyst, as well as the layman, for centuries. Despite the continuing interest, there have been few well-articulated theories of leadership formulated by social scientists. In the present review, various leadership definitions are discussed and the evolution of the concept of leadership is traced and related to current trends in leadership theory. Three broad classes of leadership theories are considered: personal ("Great," hereditary, trait); situational (*Zeitgeist*, environmental); and interactional (contingency models, path-goal, etc.). General theories of group process are discussed as they relate to the explanation of leadership. Conclusions are drawn considering the probable future of leadership theories.

Leadership has been a matter of discussion for social analysts and laymen since at least the early Greek and Roman periods. During the current century, many issues related to leadership have come under close scrutiny by investigators in several social science disciplines. In fact, a recent comprehensive review (Stogdill, 1974) cites almost 3000 studies and analyses of leadership. Despite the plethora of research and discussion, the development of theories of leadership has advanced relatively little. For the most part, it seems that social scientists have been content with empirical observation and relatively simplistic analyses based on general conceptions of what leadership is and what variables affect leadership. In the light of the development of more general theories of social interaction (e.g., Thibaut & Kelley, 1959) perhaps, as Kerr, Schriesheim, Murphy, and Stogdill (1974) say some suggest, formal theories of leadership are destined to become behavioral theory "dinosaurs" before they reach full fruition. Nevertheless, a survey of the indexes of management or psychological research will reveal a continued interest in leadership as a separate category of interpersonal dynamics. The leadership theorists are still at work, trying to bring some semblance of order to the multiplying body of research findings. The present paper is a brief review of the progress made.

A fundamental problem in leadership writings has been the lack of a generally accepted definition of leadership. The early Greeks and Romans conceived of the leader as one who began or sat in motion some activity which was subsequently completed by others (Jennings, 1960). However, the early social scientists came to think of leadership more in terms of eminence or prominence (Carlyle, 1841). This influence can be seen in Bogardus' (1934) lists of the world's great leaders and in the common man's conception of what leadership is. Some contemporary writers still seem to favor the eminence definition of leadership (Jennings, 1960), but for the most part, later observers have defined leadership in terms of social influence (e.g., Katz & Kahn, 1966), interpersonal power (e.g., Raven & French, 1958), or group process (e.g., Homans, 1950). Typically, modern definitions include reference to the group and its goals, as well as to interpersonal influence. For example, Stogdill (1950) defines leadership as "the process (act) of influencing the activities of an organized group in its efforts toward goal setting and goal achievement." Thus, in contrast to the early eminence definitions, modern definitions of leadership require a specific group setting. Bogardus (1934) lists Robert Browning and Luther Burbank as great leaders because of their literary and scientific accomplishments; the modern social scientist would want to know what groups, in what settings, were influenced by these men in the pursuit of relevant goals. In the present paper, no attempt will be made to review the various definitions of leadership. Bass (1960), Stogdill (1974), and others offer such reviews for the interested reader. The reader should keep in mind, however, that the theorist's definition of leadership can have considerable impact on his theoretical position.

A second problem in studying leadership theories is that, for the most part, what have been called "theories" in this area are more akin to orientations or empirical generalizations. Shaw and Costanzo (1970) define theory as "a set of interrelated hypotheses or propositions concerning a phenomenon or set of phenomena." As will become evident to the reader, theories of leadership do not generally fit this definition. Kerr et al. (1974) point out that most leadership theories fail to meet a number of criteria of theoretical adequacy, including: (1) non-inclusion of all relevant variables required for a full explanation of the leadership process; (2) the propositions are not a set of lawlike statements; (3) no procedures for determining the consistency of each proposition with others; and (4) many explanations for observed relationships are inadequate, and might more appropriately be considered empiric generalizations. Since a detailed critique of each theory, based on these or other criteria is beyond the scope of the present paper, the discussion to follow will center on three major themes in leadership studies and will relate these themes to our current leadership theories. The three themes are the personal, the situational, and the interactional approaches to leadership.

Personal Approaches

The emphasis in the personal approach to leadership is on the characteristics of the leader, whether inherited or learned, physical or psychological. The "Great Man Theory" of leadership is perhaps the first modern theory. Carlyle (1841) and Galton (1879) pointed social scientists toward the search for personal characteristics which could explain why some men rose to prominence while the majority remained relatively anonymous. The underlying assumption of the "Great Man Theory" is that leadership is a function of the personality and abilities of the gifted person. It emphasizes the importance of the individual. According to this approach, great men shape their environments and mold history. The research stemming from this theory is, quite naturally, the analysis of the characteristics of great leaders to see why they are great. From these analyses, various categories of "great men" have been proposed. For example, Jennings (1960) analyzes three types: (1) princes, or those who seek and maintain power; (2) heroes, or those who use power to a noble end; and (3) supermen, or those who display excellence in the midst of mediocrity.

Other personal approaches grew out of the "great man" theme. These include hereditary or biological (e.g. Woods, 1913; Wiggam, 1931) and trait (e.g., Tead, 1929) emphases. As Hollander and Julian (1969) note, the psychological study of leadership in this century began with the focus on personality and the trait approach to leadership. Bogardus (1934), though he broadened the scope of leadership study to include situational factors, devoted a great deal of energy to discussion of the characteristics which contributed to the prominence of great men--factors such as intelligence, energy, and character. Bogardus included the possibility that some leadership characteristics are learned rather than inherited. Indeed, during the early part of this century, many lists of traits the good leaders should develop were proposed.

Though the personal approaches led to many empirical studies, an increasingly confusing picture of leadership "traits" emerged. There seemed to be little congruence among various studies concerning the traits required for leadership. Bird (1940) reviewed 20 studies and found that only four of 79 traits considered were common to five or more studies. Mann (1959) reported similar results in his review. Perhaps because of this difficulty in generating an integrated picture of leadership based on personal characteristics, social scientists began to seek explanations of leadership based on situational variables.

Situational Approaches

Hemphill (1949) wrote extensively concerning the situational factors in leadership. He, among others, was looking toward a new approach to leadership, where characteristics of the group, the task, and the times are seen as determinants of what characteristics are required in a leader. According to the situational view, the person who fulfills the leadership function is the person who possesses the traits demanded by the situation. By the 1950's, as Hollander and Julian (1969) note, the situational approach had taken firm hold of the leadership field. The main advantage of the situational approach is that it readily explains why the trait theorists were unable to present a universal list of traits: the appropriate traits are a function.of the situation.

The situational approach is largely empirical rather than theoretical. Although some authors explicitly state the main principles involved, the major impact of the situational theories is in reduced concern over the personal characteristics of the leader and almost total concern with the characteristics of the situation.

It became increasingly evident, however, that the extremes of the situational approach are as detrimental to a full understanding leadership as are those of the personal approach. Clearly, both approaches maintain a misleading and artificial distinction between the leader's and the situation's influences (Hollander & Julian, 1969). Hence, after the pendulum had swung from extreme to extreme, it was almost inevitable that theories considering both personal and situational variables should appear. Indeed, such interactional theories provide the most complete and structurally satisfying theories to date.

Interactional Approaches

Most leadership theories developed within the last 20 years consider both situational and personal factors working in concert to affect group outcomes. Perhaps the best known of such interactional approaches is Fiedler's (1967) "contingency model." Citing research and reviews by other authors, Fiedler (1967, p. 4) noted that, despite numerous empirical studies, an acceptable theory of leadership had yet to be developed. As a step toward filling this void, Fiedler developed his theory of leadership effectiveness. The primary variables in Fiedler's theory are the orientation of the leader (relationship versus task), the favorability of the situation for the leader, and the effectiveness of the group. Fiedler employs paper and pencil measures of leader orientation; favorability is defined in terms of task structure and the leader's position power and affective relations within the group; effectiveness is the outcome of relevance to the group (for example, for a basketball team it would be games won). According to Fiedler, the task-oriented leader is more effective when favorability is either extremely low or extremely high. The relationship-oriented leader is more effective in moderate ranges of favorability. A large body of research, primarily conducted by Fiedler and his associates, lends the theory moderate support.

When compared to earlier theories of leadership, Fiedler's theory is normally much more acceptable. It specifies relationships between variables and yields testable hypotheses derived from these systematic relationships. However, Fiedler (1967) cautions against too broad an application of the theory, and other authors (e.g., Ashour, 1973) have criticized the theory on methodological and interpretative grounds. In spite of these limitations and criticisms, Fiedler's theory remains the standard against which other interaction theoriers are most often compared.

Another dominant influence on the interactional theories has been the research and theory stemming from the Bureau of Business Research of the Ohio State University. Based on factor analyses, two important characteristics of the leader were determined: consideration and initiation of structure. In effect, these factors separate what Fielder sees as a unitary continuum (relationship/ task orientation). According to Blake and Mouton (1964), the optimum leadership condition occurs when the leader is high in both consideration and initiation of structure. However, a number of situations have been discovered where this particular combination does not produce optimum effectiveness (cf. Korman, 1966). Hence, Kerr et al (1974) have felt the need to propose a more contingency-oriented theory based on consideration and initiation of structure.

Evans (1970) and House (1971) propose a Path-Goal Theory of Leadership based on the Ohio State Leadership Studies. According to the Path-Goal Theory, the leader's consideration determines the followers' perceptions of the rewards available, while his initiation of structure determines the followers' perceptions of the paths by which the rewards may be obtained. The leader's main functions are to increase the desirability of the rewards for his followers and to clarity path-goal relationships. House (1971) reviewed the consideration/ initiation of structure literature in support of the theory, and he offers several hypotheses derived from his review and from the theory. As yet the theory remains largely untested.

Numerous other theories of leadership have been proposed, representing various theoretical perspectives. For example, Bass (1960) proposes a reward/ punishment theory, and Likert (1961) offers a human-relations orientation. Though these, and other, authors have contributed somewhat to our understanding of leadership, each of their approaches fails to qualify as an acceptable general theory of leadership. Of course, there are many theories of group dynamics which explain leadership as one aspect of group interaction (e.g., Thibaut & Kelley, 1959). A review of such theories is beyond the scope of this paper.

Conclusion

In tracing the development of leadership theory during this century, one might gather the impression that the progression of leadership theories has been orderly. This is not the case; the flow has been neither smooth nor regular. Despite the many empirical investigations and theoretical interpretations, we remain without an adequate formal theory of leadership. But, the future is promising. Fiedler (1967)

and Kerr et al. (1974) point the way toward a more systematic leadership theory, one which can be developed and tested, and improved. Time will tell whether these beginnings will culminate in a full theory of leadership. One would certainly hope that they do.

REFERENCES

Ashour, A.S. The contingency model of leadership effectiveness: An evaluation. *Organizational Behavior and Human Performance,* 1973, *9,* 339-355.

Bass, B.M. *Leadership, phychology, and organizational behavior.* New York: Harper, 1960.

Bird, C. *Social Psychology.* New York: Appleton-Century, 1940.

Blake, R.R., & Mouson, J.S. *The managerial grid.* Houston: Gulf, 1964.

Bogardus, E.S. *Leaders and leadership.* New York: Appleton-Century, 1934.

Carlyle, T. *Heroes and hero worship.* Boston: Adams, 1907 (original 1841).

Evans, M.G. The effects of supervisory behavior on the path-goal relationship. *Organizational Behavior and Human Performance,* 1970, *5,* 177-198.

Fiedler, F.E. *A theory of leadership effectiveness.* New York: McGraw-Hill, 1967.

Galton, F. *Hereditary genius.* New Yark: Macmillan, 1914 (original 1879).

Hemphill, J.K. *Situational factors in leadership.* Columbus: Ohio State University, 1949,

Hollander, E.P., & Julian, J.W. Contemporary trends in the analysis of leadership processes. *Psychological Bulletin,* 1969, *71,* 387-397.

Homans, G.C. *The human group.* New York: Harcourt, Brace, 1950.

House, R.J. A path goal theory of leader effectiveness. *Administrative Science Quarterly,* 1971, *16,* 321-338.

Jennings, E.E. *An anatomy of leadership.* New York: Harper, 1960.

Katz, D., & Kahn, R.L. *The social psychology of organizations.* New York: Wiley, 1966.

Kerr, S., Schriesheim, C.A., Murphy, C.J., & Stogdill, R.M. Toward a contingency theory of leadership based upon the consideration and initiation of structure literature. *Organizational Behavior and Human Performance,* 1974, *12,* 62-82.

Korman, A.K. "Consideration," "initiating structure," and organizational criteria—a review. *Personnel Psychology,* 1966, 19, 349-361.

Likert, R. *The human organization.* New York: McGraw-Hill, 1961.

Mann, R.D. A review of the relationships between personality and performance in small groups. *Psychological Bulletin,* 1959, *56,* 241-170.

Raven, B.H., & French, J.R.P. Legitimate power, coercive power, and observability in social influence. *Sociometry.* 1958, *21,* 83-97.

Shaw, M.E., & Costanzo, P.R. *Theories of social psychology.* New York: McGraw-Hill, 1970.

Stogdill, R.M. Leadership, membership, and organization. *Psychological Bulletin,* 1950, *47,* 1-14.

Stogdill, R.M. *Handbook of leadership.* New York: Free Press, 1974.

Tead, O. The technique of creative leadership. In *Human nature and management,* New York: McGrawHill, 1929.

Thibaut, J.W., & Kelley, H.H. *The social psychology of groups.* New York: Wiley, 1959.

Wiggam, A.E. The biology of leadership. In H.C. Metcalf (Ed.), *Business leadership.* New York: Pitman, 1931.

Woods, F.A. *The influence of monarchs.* New York: Macmillan, 1913.

DEVELOPMENT AND EVALUATION OF ARMORED CAVALRY ENGAGEMENT SIMULATION

Claramace S. Knerr, Earl S. Stein, and Bruce W. Hamill
US Army Research Institute for the
Behavioral and Social Sciences

Maj Angelo A. Severino
US Army Training Support Center

Engagement simulation (ES), designed for high psychological fidelity in Army tactical combined arms training, is a recently implemented training approach. Emphasis is placed on realistic field training situations and procedures (Root, 1976). ES is conducted using two-sided, free-play exercises where outcomes are not preplanned, but depend on the behavior of teams and individuals. The training field, weapons, procedures, and associated troops form a stimulus system. Individual or team actions serve as potential stimuli so that the situation evolves based on what people are actually doing rather than on a fixed scenario. Thus, the situation is emergent rather than prespecified, highly predictable, or amenable to analytic solution (Boguslaw and Porter, 1966). As the action unfolds, it provides performance training under realistic tactical conditions in a discovery, or trial and error, paradigm, coupled with structured feedback.

All weapons effects and signatures are simulated to represent the real battle sounds, sights, and smells. Through the use of pyrotechnics and blank adaptors, the flavor of battle is captured without its inherent hazards. Casualties are assessed using rules of engagement and decision processes which are based on actual hit kill probabilities and have face validity for the participants.

Early ES concentrated on rifle squads. Casualties were inflicted by soldiers noting helmet numbers of opposing team members and yelling the sighted numbers to control personnel who radioed the "kill" to opposing team controllers (US Army Infantry School, 1975). Control personnel, who apply objective rules of engagement and casualty assessment, are a critical component of the ES training system. They supply immediate and definite engagement feedback by reporting and confirming the "kills."

Each ES exercise is followed by an After Action Review (AAR) which recreates the action and provides additional information to the soldiers as to the consequences of their behavior. Soldiers from both opposing forces participate in the AAR so that feedback from the "enemy" reinforces learning through peer diaglogue.

The discussion leader for the AAR is a senior controller. He coordinates the controllers, directs the exercise as a whole, and acts in the role of unit commander. The AAR leader uses a record of the casualties to guide the discussion, which recaps the exercise chronologically. The record is kept in an exercise Net Control Station (NCS) where personnel write the time and note the tactical elements involved in engagements. Thus, the sequence of casualties is recorded for the AAR leader to guide the discussion.

ES has been implemented for infantry, armor, and antiarmor employing platoon sized units with combined arms force mixes. Known as REALTRAIN, it has been validated in Europe (Root, Epstein, Steinheiser, Hayes, Wood, Sulzen, Burgess, Mirabella, Erwin, & Johnson, 1976). The present report concerns the development of ES procedures for armored cavalry units.

"Cavalry's basic tasks are reconnaissance and security" (Department of the Army, 1975), so that the armored cavalry platoon performs information gathering and reporting functions. Reconnaissance missions often do not lead to the casualty producing engagements typical of other ES training. In some instances, they are one-sided, so that no engagements occur, such as reconnaissance of an area that does not contain enemy elements. In other instances, enemy elements are present, so that the exercise is two-sided. If the opposing forces fire, then the exercise becomes a standard casualty producing one. In contrast, if they withhold fire, reconnaissance functions are performed, and given the appropriate procedures, can be reenacted in the AAR. The problem was to develop and field test techniques to stimulate reconnaissance functions in ES exercises, while maintaining the objectivity and credibility of the casualty producing exercises. In addition, appropriate emphasis had to be placed on reconnaissance in the AAR so that the desired behavior would be supported and maintained.

METHOD

Procedures were drafted for armored cavalry ES, including the control system, weapons effects and signature simulators, rules of engagement, casualty assessment, and special techniques for incorporatnng the reconnaissance functions into the exerises and AAR. The procedures were examined and refined in a series of field tests. These tests were developmental in nature, rather than validations.

After some small scale exploratory tests, the procedures were examined in platoon against platoon exercises, conducted in May and November 1977, with troop support provided by the 3d Armored Cavalry Regiment, Ft Bliss, Texas. Following preliminary controller training, six days of exercises were run with one exercise and AAR each day. The opposing forces were two armored cavalry platoons in the first test, and were one cavalry platoon and one combined arms team in the second test. Up to 90 soldiers participated in each exercise.

Data collection instruments were designed to process objective and subjective data. Objective data included the NCS casualty record, maps and notes made by the senior controllers and element controllers, and tabulation forms for recording reconnaissance activities. Questionnaires were written to assess subjective judgments of the participants and controllers regarding training value, simulator credibility, casualty assessment, utility of the procedures, and various aspects of the AAR.

RESULTS AND DISCUSSION

The first test demonstrated the effectiveness and utility of the weapons effects and signature simulators, rules of engagement, and casualty assessment. Improvements in controller optics for some weapons, and in the mortar procedures, were made on the basis of this test, as described in detail elsewhere (Knerr, Severino, & Bosley, 1977).

ES training motivates the participants to close with and destroy the opposing force. This traditional casualty oriented behavior is counterproductive when the goal is to obtain information rather than become decisively engaged. Reconnaissance had been minimal in the earlier test due to the motivating contingencies of standard ES exercises, where knowledge of results is rapid. Feedback for information gathering

and reporting did not occur until the AAR. To emphasize reconnaissance, several interrelated techniques were employed: exercise scenarios were designed to maximize information and reporting. The activities of control personnel supported this goal by their actions with the trainees, the detailed records they kept, and their discussion of reconnaissance during the AAR.

The effectiveness of the exercise scenario was evident from the start of the first exercise, and continued throughout the test. The opposing forces had missions requiring information gathering and reporting (zone reconnaissance for the cavalry platoon, delay for the combined arms team). The tactical situations described in the operations orders were realistic and capable of sustaining a weapons hold situation, where direct fire would only be used with the permission of the unit commander, who was the senior controller for each force. The vehicle commanders reported large amounts of information concerning the enemy, along with repeated requests for release from weapons hold, and consequent permission for direct fire. They used their reports to build credibility and convince the senior controller of their need to fire. Thus, the weapons hold status, applied in the highly motivating ES enviornment, appeared to elicit concentrated reconnaissance activity. Permission to fire appeared to become a reward for building a credible reconnaissance reporting record.

Three methods of tactical information storage were employed during the exercises. These were the NCS sheet, 3x5 cards, and map sketches. These records were maintained by the senior controllers, or by assistants who accompanied them. The official NCS record was maintained in NCS in all exercises, but in some exercises, the senior controller recorded tactical notes on an NCS form for his use during the AAR. Each type of record became a potential data base for the AAR. Success of each type of information record depended primarily on the idiosyncratic preferences of the senior controller leading the AAR. Each preferred the record he kept himself regardless of which type it was and irrespective of any convenience that had been predicted.for it.

The senior controllers evolved a system of coordination so that each knew where all personnel were and what the tactical effects would be if they gave any element permission to fire. The communication that evolved between the senior controller, who would be the AAR leader, and his counterpart with the opposing team was initiated to control the weapons hold status and the behaviors which it generated. This communication became a vehicle to keep the senior controller informed during the exercise such that he was apparently able to integrate the tactical information conceptually and easily apply it during the AAR discussion.

Controllers assigned to tactical elements had special duties related to reconnaissance functions. They were thoroughly briefed on the exercise scenario, operations orders, and expected tactical events before the exercise. During the action, they recorded tactical events on maps and 3x5 cards to enable verification of tactical activities. Maps were distributed to a total of 37 controllers over two days of training, and were returned by 35 of them with at least notes from the operations orders. Twenty maps (i.e., by 54.1% of the controllers) contained notes and location sketches in addition to the operations order information, and three controllers also wrote the ES numbers for vehicles and individuals on the maps. Controllers sketched maps to the same extent regardless of the type of element they controlled, with the exception of the infantry controllers. Only one infantry controller in ten possible instances wrote notes or sketched on the map. Infantry controller map usage was much lower than that of other controllers.

Usage rates of the 3x5 cards were very high during all six exercises: 97 of the 113 total possible cards were used to record ES numbers or notes or both (95 cards has ES numbers, and 60 cards had tactical notes). The controllers recorded a high volume of tactical notes daily, although they were not asked to keep the records every day (25-30% of the controllers made notes on days when they were not explicitly asked to do so). Infantry controllers used the cards to the same extent as did other controllers. Thus, the 3x5 cards proved feasible for infantry, in contrast to the maps.

Participants and controllers were asked for their subjective evaluations of the training value of the exercise and how ES exercises compared with other training. Participants (N = 77) responded as follows to the question "How much would you say you learned during the training exercises you have just completed?": A great deal, 44%; Some, 38%; Little or nothing, 18 %. When asked to compare the ES exercises to other training, most participants replied that the ES exercises were better: REALTRAIN much better, 36%; REALTRAIN better, 43%; No difference, 11%; REALTRAIN worse, 11%. Compared to the REALTRAIN same portion of the questionnaire answers are in the combined "better" and "much better" categories, but in the Europe data, the majority responded that the ES exercises were "much more effective." It should be emphasized that the armored cavalry exercises entailed development of a new system, in contrast to the Europe validation of smoothly conducted training.

Controllers (N = 48) were asked how much they learned about tactics during the exercises. Responses showed that they perceived that they were learning, often as much or more than if they were part of the tactical team: Learned more as a controller, 54%; Learned a fair amount as a controller, 33%; Didn't learn much about tactics when controlling, 13%.

Summary

During the last five years, Engagement Simulation has come of age as the U.S. Army's effort to develop a two-sided, free-play, situation dependent (emergent) tactical training experience for combat soldiers. Employing weapons signature simulation and accurate casualty assessment produces a psychological climate where fidelty is as high as it could be, short of actual battle. Training occurs based on experiential learning reinforced by immediate feedback through success or failure and through delayed feedback via an After Action Review.

REALTRAIN for Armored cavalry faced an additional problem brought about by the high level of motivation generated by using weapons systems and receiving rapid feedback of success. Trainees lost sight of the importance of reconnaissance. A control system was evolved based on the learning principle where high probability behavior (firing at the opposition) was used to reinforce behavior of lower probability (providing reconnaissance information) (Premack, 1965). This proved to be very satisfactory for increasing the number of reports which were sent from small unit leaders to their commanders.

A secondary problem in sustaining reconnaissance was the storage and accurate retrieval of tactical information for the After Action Review. Encoding that information in the form of notes appeard to be a very individual process and a formal structure for note taking was less relevant to retrieval than the opportunity of the discussion leader to use his *own* notes regardless of their format.

REFERENCES

Boguslaw, Robert, & Porter, Elias H. Team functions and training. In R.M. Gagne and A.W. Melton (eds), Psychological principles in system development. New York: Holt, Rinehart, and Winston, 1966.

Department of the Army. Field Manual No. 17-95, *Cavalry*. Washington, DC, 1 July 1977.

Knerr, Claramae S., Severino, Angelo A., & Bosley, John J. *Initial test of Armored Cavalry engagement simulation*. Paper presented to the annual meeting of the Military Testing Association, San Antonio, TX, October, 1977.

Premack, A. Reinforcement Theory In M.R. Jones (Ed) *Nebraska Symposium on Motivation*, Lincoln: University of Nebraska Press, 1965.

Root, R.T. *Engagement simulation: A training environment for the learning of complex tactical skills*. Paper presented to the Military Operations Research Society, Ft Eustis, VA, December, 1976.

Root, R.T., Epstein, K.I., Steinheiser, F.H., Hayes, J.F., Wood, S.E., Sulzen, R.H., Burgess, G.G., Mirabella, A., Erwin, D.E., Johnson, E. *Initial validation of REALTRAIN with Army combat units in Europe*. ARI Research Report 1191, October 1976.

U.S. Army Infantry School, *Squad combat operations exercise simulation* (SCOPES), Training Circular St7-2-172, 1975.

COGNITIVE PRETRAINING: AN AID IN THE TRANSITION FROM INSTRUMENT TO COMPOSITE FLYING

John V. Crosby
Department of Psychology
Arizona State University

Abstract

The present study was designed to investigate the utility of cognitive pretraining relative to the early difficulties encountered by student pilots in transition from ground based instrument training to composite flying training. The cognitive pretraining consisted of: (a) an instrument reading review, (b) a vocabulary of relevant cockpit features, (c) the use of brief perceptual rules for pitch and bank attitudes, and (d) prototype representations of a variety of pitch and bank attitudes. Three groups of twelve pilots each participated in the study: student experimental, student control, and experienced instructor pilots (IPs). The experimental group was exposed to cognitive pretraining and then compared to the student control and IP groups in a simulated composite flight laboratory task. Results of the laboratory task demonstrated superior discrimination performance of the student experimental group over both the student control and experienced pilot groups for the most difficult discrimination. As the discrimination difficulty decreased, the performance of the experimental and experienced pilot groups became equal and both were superior to the student control group. No between group differences in RT were detected. As a measure of the external validity of the laboratory task, both student groups were subjected to four discrete aircraft maneuvers in the Williams Air Force Base Human Resources Laboratory (HRL) Advanced Simulator for Pilot Training (ASPT). Results of the ASPT task support the findings of the laboratory task.

The U.S. Air Force Undergraduate Pilot Training (UPT) program at Williams Air Force Base is designed to provide student pilots with the academic and perceptual-motor skills necessary for the safe and proficient flying of high performance jet aircraft. Central to the concept of UPT is the concern that student pilots acquire and further develop good judgment skills rather than fixed stimulus-response connections. Good judgment skills can best be described in terms of adaptability and flexibility of performance based upon a variety of situational factors rather than a fixed network of generalizable responses. Wood (1973) in a survey of relevant learning hypotheses, suggested several possible dimensions influencing flying performance. These dimensions contain both cognitive and motor attributes and range from a clear understanding of the principles of the task to the possession of a motor program of the required responses.

In the early phases of UPT, student pilots receive instructions and task demonstrations together with opportunities for supervised practice in ground based simulators (ATC Syllabus, 1976). Prior to the actual flying or composite phase of training, student pilots receive instruction on complex emergency procedures, cockpit checks, and instrument maneuvers such as straight-and-level, constant-rate-climb, turn-to-heading, etc. It is during the instrument phase of pilot training that the students' initial cognitive and perceptual-motor skills are developed. During this early phase of pilot training, the student pilot receives knowledge of results regarding task performance exclusively from the instruments. That is, the student pilot attends to and subsequently receives visual feedback from an array of aircraft instruments in what could essentially be described as a tracking task. Once in the aircraft or flying phase of pilot training, however, the student pilot is expected to function in a composite mode and as such divide his attention in an 80:20 fashion; i.e., spend 80 percent of his time attending to information external to the cockpit and 20 percent of his time monitoring. The instruments (ACTM, 1974; T-37 Instructor Techniques, 1974). (Composite flight is a flying term used to describe the technique of using external references, supported by flight instruments, to establish and maintain aircraft flight attitude, ACTM *51-38, 1973.*)

Instructor Pilots (IPs) generally report that student pilots experience difficulty in the transition from basic instrument to composite flying. This difficulty is believed to result from the student pilots' early dependency on the instruments as the predominant source of visual feedback and the absence of a visual representation of the "view-from-the-cockpit." This new requirement of having to spend 80 percent of his time attending to information external to the aircraft and only 20 percent of his time monitoring the instruments is in direct conflict with his earlier training. Assuming these observations to be correct, the student pilot transitioning from the instrument, or ground based simulation phase of training, to the composite flying phase of training must somehow reorder his source of feedback information regarding the attitude of the aircraft from that solely obtained from instruments to a combined or integrated source composed of both external and instrument information.

If transition difficulties are due primarily to early dependence on a particular source of feedback, i.e., instruments, what effect would cognitive pretraining on external visual references have on the transition from instrument to composite flying? Would cognitive pretraining facilitate the transition by providing early conceptual information regarding the attitude of the aircraft as viewed from the cockpit and therefore assist the integration of the two sources of information? Or would such pretraining be irrelevant?

This research is designed to probe those questions. Specifically, what is the role of cognitive pretraining as an aid in the transition from basic instrument to composite flying?

METHOD

Subjects

The experimental and control groups consisted of twelve U.S.A.F. Undergraduate Pilot Training (UPT) student pilots, drawn from two Williams Air Force Base UPT classes. Six of the twelve students from each class were randomly assigned to the experimental group with the remaining six assigned to the control group. The conditions prerequisite for participation in the study were that the student pilots have less than 50 hours of flying time and no previous rating (e.g., navigator). Also, foreign students were excluded. An additional, external control group, consisting of twelve Instructor Pilots (IPs) provided experienced pilot data.

Pretraining Materials

The cognitive pretraining material consisted of: (1) a review of instrument reading; (2) a description and feature vocabulary of the various canopy references, i.e., canopy bow, windscreen, glare shield, center bow, etc.; (3) the identification of the canopy references together

with the vertical and horizontal reference planes depicting the conceptual nose of the aircraft, and (4) the aircraft attitude relationship with reference to the horizon for straight and level, left turn, right turn, nose high, and nose low conditions. All pretraining materials were developed from operational flying training information extracted from a variety of sources at Williams Air Force Base and independently verified as accurate representations by research and flight line IPs prior to use. Achievement tests were administered following self-paced instruction and prior to laboratory testing to assess the degree of learning. No *a priori* learning criterion was used. All pretraining materials were collected prior to the experimental task.

Laboratory Task, Apparatus, and Stimuli

To assess experimentally the student pilots' grasp of the concepts developed in the cognitive pretraining phase, a laboratory task consisting of tachistoscopically presented instrument and external horizon slides was used. First, a colored slide of the instrument panel was briefly displayed. The slide contained all information necessary to determine the attitude of the aircraft, i.e., nose high, nose low, left, or right turn. Following the instrument slide, a colored slide depicting the aircraft's attitude (as viewed from the cockpit) was briefly presented. Pitch and bank conditions were counterbalanced for order effects. The task of the subject was to determine whether the second slide was correct or incorrect based on the attitude information contained in the first slide. The reverse conditions by the instrument slide was also run to assess any performance differences due to presentation sequence.

The stimuli consisted of two broad attitude categories: (1) straight climbing and diving pitch changes for the pitch condition, and (2) left and right level turns for the bank conditions. Seven stimulus conditions were included within each of the two broad categories with the *a priori* probability of being correct being constant at 50 percent. That is, for each of the seven stimulus value conditions there were six correct and six incorrect slides. For example, if the instrument slide depicted wings level 0° pitch, six of the contact slides were wings level 0° pitch while the remaining six described $+2^\circ$, $+4^\circ$, $+6^\circ$, and -2°, -4°, -6° nose high and nose low conditions respectively.

Procedure

Student pilots were individually scheduled to minimize negative impact on any phase of their training program. Practice slides were presented to the subject and sample responses were requested to verify understanding of the task. A response of correct indicated that the second slide correctly represented the attitude of the aircraft based on information contained within the first slide. Once initiated the stimulus presentations were cycled automatically. Two dependent measures were taken, correct/incorrect and response latency.

Measure of External Validity

The USAF/HRL Advanced Simulator for Pilot Training (ASPT), a high fidelity replication of the T-37B Jet Trainer, was utilized to obtain a measure of the external validity of the pretraining. ASPT is completely described by Bell (1974) and will not be discussed here. Both the experimental and control students were required to fly three trials of four maneuvers in each of two visual conditions: (1) instruments only, and (2) composite reference to the horizon with the attitude indicator covered. A rated IP was present in the cockpit and provided basic aircraft control related information before and after, but not during maneuvers. Computer generated performance measurement data in the form of RMS error scores for altitude, airspeed, heading, and bank angle were taken. Additionally, at the completion of each discrete trial, the IP provided a single global score for the maneuver on a scale of 1 to 12, with 1 unsatisfactory and 12 excellent.

RESULTS

Several analyses were performed for both the experimental laboratory task and the HRL/ASPT simulation task. The results are reported under appropriate subheadings in order to facilitate communication of the data.

Laboratory Task

It was expected that the performance of the student pilot group receiving the cognitive pretraining would be superior to the student control group in the experimental laboratory task. To assess the between group differences, percent correct responses were computed for each of the three levels of discrimination difficulty for what was believed to be the two most representative stimulus conditions, i.e., 0° Pitch and 0° Bank. These data, presented graphically in Figure 1 and 2 were subjected to separate 3 (Group) X 2 (Sequence) X 3 (Difficulty) repeated measures analysis of variance.

For pitch, significant main effects were observed for Group, $F(2,33) = 5.233$, $p < .02$ and Difficulty, $F(2,66) = 74.234$, $p < .001$. For Bank, significant main effects were found for Group, $F(2,33) = 10.700$, $p < .001$ and Difficulty, $F(2,66) = 32.770$, $p < .001$. Additionally, in the bank condition, a Group X Difficulty interaction was observed, $F(4,66) = 10.453$, $p < .001$ as well as a marginal Group X Sequence X Difficulty interaction, $F(4,66) = 2.870$, $p < .05$.

A single 3 (Group) X 2 (Attitude) X 2 (Sequence) repeated measures analysis of variance of response latency (RT) indicates no significant between group differences, $F(2,33) = .833$, $p = .553$. A marginal main effect for Attitude, $F(1,33) = 3.93$, $p = .05$ indicated the RT to bank attitude was faster than RT to pitch attitude and a Sequence main effect indicates a reliably faster RT for the contact/instrument sequence, compared with the instrument/contact sequence, $F(1,33) = 21.66$, $P < .001$.

In summary, the results of the laboratory task analyses indicate that: (a) the experimental group (for the 0° Pitch and 0° Bank conditions) performed superior to both the student control and the experienced IP groups at the more difficult discrimination levels, (b) as the degree of discrimination difficulty decreased, both the experimental and IP groups maintained reliable superiority over the control group but did not differ significantly between themselves, and (c) no significant between group differences in RT were detected although within group differences for Attitude and Sequence were present.

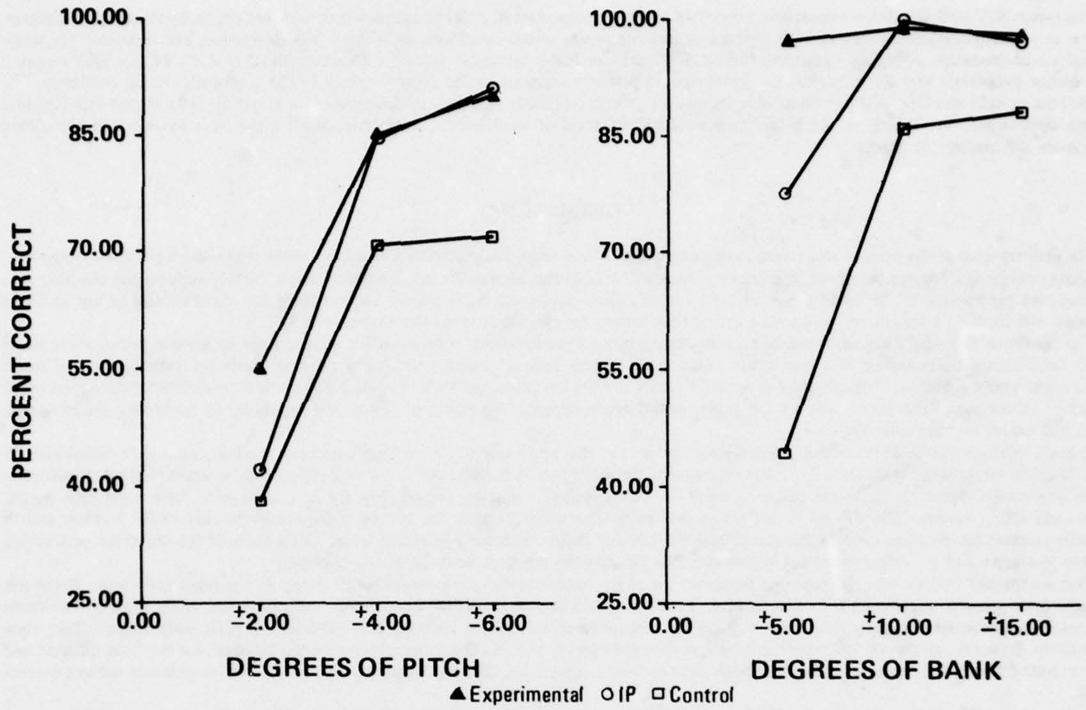


FIGURE LEGEND

Figure 1. Percent correct pitch discrimination as a function of discrimination ease.

Figure 2. Percent correct bank discrimination as a function of discrimination ease.

Simulation Task

To gain some insight into the external validity of the cognitive pretraining and of the laboratory task, three trials of four representative aircraft maneuvers in each of two visual conditions were flown in ASPT. The maneuvers selected as being the most representative of the student's skill repertoire for their level of training were: (1) vertical-s-alpha, (2) straight-and-level, (3) turn-to-heading, and (4) steep-turn. Of the four maneuvers selected, all but the vertical-s-alpha, had been practiced repeatedly in the ground based instrument simulator and all students had satisfied the maneuver proficiency requirements set forth in the instrument phase of training. For all ASPT maneuvers, multiple dependent measures in terms of RMS error scores and IP ratings were taken. Altitude, Airspeed, and IP ratings were constant dependent measures taken across all maneuvers. Heading scores were taken for all but the steep-turn and turn-to-heading maneuvers. Climb and descent Rate measures were taken for only the vertical-s-alph maneuver.

Four separate 2 (Group) X 2 (Visual Condition) X 3 (Trials) repeated measures multivariate analyses of variance were computed. Generally, only significant multivariate results are reported unless specific *a priori* hypotheses predict given dependent variable effects (Finn, 1974).

In the analysis of the vertical-s-alpha maneuver, the experimental group performed significantly better than the control group, $F(5,18)=4.513$, $p<.01$. It will be recalled that the vertical-s-alpha was the only maneuver of the four maneuvers flown that was novel to both groups. There were no significant between-group differences for the remaining three maneuvers.

A main effect for Visual Condition for all four maneuvers indicated superior instrument performance over composite performance for both groups. The multivariate results are: for vertical-s-alpha, $F(5,18)=13.057$, $p<.001$; for straight-and-level, $F(4,19)=7.877$, $p<.001$; for turn-to-heading, $F(5,18)=6.993$, $p<.001$; and finally, steep-turn, $F(4,19)=9.220$, $p<.001$. A main effect for Trials was present for straight-and-level, $F(8,15)=3.143$, $p<.05$ and for turn-to-heading, $F(10,13)=2.833$, $p<.05$. These findings indicate a significant practice effect. No additional significant multivariate results were detected.

It was hypothesized that the experimental group would perform superior to the control group when compared across specific discrete dependent measures. For example, as a result of cognitive pretraining, it was believed that the experimental group would perform superior *to the control group in Heading, Bank, and Rate. In light of these a priori hypotheses, the most important* analyses are the univariate values within the Group main effects and the Group X Visual Condition interactions. Inspection of the univariate values for the vertical-s-alpha maneuver indicates reliable findings supporting the superior performance of the experimental group over the control group for Rate, $F(1,22)=15.869$, $p<.001$; Heading, $F(1,22)=10.498$, $p<.005$, and to a lesser extent, Airspeed, $F(1,22)=6.207$, $p<.025$.

The group X Visual Condition interaction univariate terms are important in order to assess whether or not the between groups difference is due to performance differences in both instrument and composite visual conditions or whether the differences are restricted primarily to one visual condition. A highly significant Group X Visual Condition univariate value for Heading, $F(1,22)=10.124$, $p<.005$ supports our earlier hypothesis that the experimental group would perform superior to the control group in the composite visual condition. For each of the straight-and-level and turn-to-heading maneuvers, the univariate between group differences for Heading, favoring the experimental group, approached, but failed, to reach the traditional $p<.05$ level of significance. Inspection of all maneuvers revealed no significant univariate differences for Bank.

DISCUSSION

The primary goal of the present study was to determine the role of cognitive pretraining as an aid in the transition from basic instrument to composite flying. The results of the pretraining, assessed by both the laboratory and simulation tasks, clearly support the effective role of cognitive pretraining in the early phase of pilot training. The improved discrimination accuracy of the experimental group in terms of pitch and bank was supported by parallel differences within the discrete maneuvers flown in ASPT.

The results of the ASPT simulation task clearly demonstrates a performance decrement for *both* groups of student pilots when asked to fly familiar maneuvers while in a composite visual mode. These findings indicate that early training solely on instruments may result in the student pilot acquiring instrument tracking skills rather than an integrated network of skills based on internal (instruments) and external (horizon) information. The superiority of the experimental group suggests the utility of cognitive pretraining in facilitating the transition from instrument to composite flight.

Overall, perhaps the most powerful effect demonstrated was the improved ability in discriminating shallow, i.e., $\pm 5^{0}$ differences in bank from the wings level conditions. This was indicated by the difference in heading performance between the two student groups. Instructor pilots universally report that students naive to the T-37 flying phase of training consistently fly in a shallow 5^{0} bank right turn during their early flying sorties. The IPs interpret this as the students' attempt to align the curved glare shield parallel to the horizon and to visually position the physical nose of the aircraft on the horizon rather than the conceptual nose. The effects of the cognitive pretraining appear to clarify and provide conceptual references that tie into the physical features of the cockpit.

One unexpected finding was the superior performance of the experimental group over the IP group in the laboratory task. There are perhaps three possible explanations for this finding. The first and most obvious is that the stimulus materials used, while initially confirmed as accurate representations, may have been sufficiently inaccurate to result in a decrement in experienced pilot performance. This view is unlikely, however, as the student experimental and the experienced pilot accuracy data were essentially equal for the least difficult and next to least difficult discrimination levels for both pitch and bank conditions. The control group leveled off at a much lower percent correct level.

A second possible explanation for the difference may lie in the observations that the IP normally flies in the right seat of the aircraft and therefore may utilize different physical cockpit features when making attitude discriminations. Also, the IP does not have an attitude indicator and must glance across the cockpit to observe the student's attitude indicator. This is also unlikely as performance should have been suppressed across all levels of difficulty.

A third possible explanation is that the experienced pilots, while once heavily dependent on external visual cues anchored within the physical features of the cockpit, have now transcended that cue dependence and utilize a more integrated visual and kinesthetic based feedback repertoire. Further research is needed to evaluate this assumption.

REFERENCES

Bell, F.E., III *Advanced Simulation in Undergraduate Pilot Training (ASUPT)*. Technical Fact Sheet, Air Force Human Resources Laboratory Technical Report: AFHRL/FT TN 73-01, May 1974.

Finn, J.D. *A general model for multivariate analysis*. New York: Holt, Rinehart, & Winston, Inc., 1974.

United States Air Force. *Advanced flying: Jet* (Air Training Command Manual 51-38). Washington, DC: Department of Defense, 1973.

United States Air Force. *Primary flying: Jet* (Air Training Command Manual 51-4). Washington, DC: Department of Defense, 1974.

United States Air Force. *T-37 instructor techniques* (Randolph AFB Study Guide F-V5A-A-IT-SG). Randolph Air Force Base, Texas: Department of the Air Force, 1974.

United States Air Force. *Syllabus of instruction for undergraduate pilot training* (T-37/T-38) (Air Training Command Syllabus P-V4A-A). Washington, DC: Department of Defense, 1975.

Wood, M.E. Transfer from audiovisual pretraining to an instrument flight task. Unpublished doctoral dissertation, Arizona State University, 1973.

DEVELOPMENT OF A NONELECTRONICS MAINTENANCE TRAINING SIMULATOR

John F. Brock
Navy Personnel Research and Development Center

Although simulators for training have come into their own, their use to train system maintenance tasks has been limited. Even more limited has been the application of simulator technology to nonelectronics maintenance training. This paper reports on an instructional system design (ISD) effort which produced both hardware and software specifications for a minicomputer-based simulator which is used to train preventive and corrective maintenance of a pneumatic control system. The specific techniques for deriving the simulation specifications are discussed as well as the use of ISD to develop more conventional instructional media.

INTRODUCTION

The past decade has seen all three services increasingly interested in some sort of systematic approach to the design of instruction. However, as Montemerlo and Tennyson (1976) point out, there is no established, generally accepted process by which instruction can be designed. However, as system analytic techniques have been applied to instructional development, some general principles of design have evolved (e.g., Chief of Naval Education and Training, 1975).

A subset of the larger problem of designing instruction is selecting appropriate media to conduct instruction. Again, no single media selection model or guide is presently available. Most current models either stress a cost-benefit approach (e.g., Braby, Henry, Parrish, and Swope, 1975) or an approach linking media to some kind of behavioral or learning type (e.g., Merrill, 1971). Only recently have attempts been made to link skills to media to costs (e.g., Stewart and Kelly, 1977).

One instructional technique which has attracted a lot of attention and R&D money is simulation. Maintenance training simulation has only recently received some of this attention, particularly in the Air Force (e.g., Miller and Rockway, 1975) and the Navy (Brock, 1977a). However, nearly all of these simulation studies have been concerned with training electronics and avionics maintenance tasks.

The Navy has made some progress toward applying the systems analytic approach to training design and the maintenance training simulation technology to a nonelectronics maintenance system. The system selected for this initial effort is the Hagan Automatic Boiler Control (ABC) system—a kind of carburetor for the main propulsion plants of many Navy ships (for a more complete description of this system and why it was chosen, see Malone, DeLong, Farris, and Krumm, 1976). The system is maintained by mid-level enlisted personnel (E-5, E-6) who must troubleshoot the system, isolate malfunctions within it, disassemble and repair components, and assemble and calibrate components. The system itself is pneumatic; it delivers signals to a main console via air pressure changes introduced by main propulsion system sensors located throughout the length and breadth of the propulsion plant. The console contains personnel and the regulatory components which send out air pressure signals to various remote control mechanisms. This console is the primary interface which maintenance persons have with the system. It is this console which is simulated and which will discussed below.

The purpose of this paper is to briefly describe the instructional design process which was used to produce the new Hagan ABC maintenance course and to report the results of the pilot test of that new course.

Instructional Design Prodedure

A complete description of the instructional development process is reported in Malone et al. (1976). The basic approach stems from work done by the author and a group of Navy Chief Petty Officers at the Propulsion Engineering School, Great Lakes, Illinois (Brock and DeLong, 1975). Essentially, the process takes the course developer from a set of logically derived job tasks through the development of behavioral objectives to the conduct of a job relevant training course. The key element in this process is the identification of those skills which must be acquired if the job is to be accomplished.

In the present case, this kind of analysis produced a set of cognitive skills which had not been previously identified for Hagan ABC maintenance personnel; these were the logic systems the maintenance man must use to troubleshoot the Hagan system. A set of algorithms was developed which have as their start point a symptom (e.g., black smoke is coming from the ship's stack) and, as their end point, a diagnosis (the XYZ valve is stuck) and an action (remove and repair the XYZ valve). Although algorithmic sequences abound in system maintenance activities, to the writer's knowledge, this was the first attempt to make major instructional decisions for an entire training course based upon them. Previous efforts have concentrated on smaller units of instruction (e.g., Miller and Rockway, 1975).

Because of the algorithmic nature of the troubleshooting tasks, some computer-based instruction seemed appropriate. However, small student loading (15 students per class; six classes per year) was a persuasive argument against a full-blown computer assisted instruction system. Some mix of conventional instructional techniques with low-cost computer usage was deemed more reasonable.

Instruction Design Results

The final product is a self-paced training program which provides opportunities for hands-on training. Instructional delivery is accomplished with handouts, workbooks, and the Bessler Cue-See audiovisual device. Troubleshooting instruction is conducted on a computer-based simulation system which allows for skill acquisition, student practice, and testing. Since it is the most innovative aspect of the system, it will be discussed in some detail below.

However, some discussion of how the conventional media were selected is needed. The contractor, Essex Corporation, used the method outlined in Figure 1. Space limits a discussion of this technique, but as the writer has pointed out elsewhere (Brock, 1977b), the problem of selecting the best instructional medium to maximize the acquisition of some specific skill or knowledge is remarkably intractable. As neat and clean as the process in Figure 1 appears on paper, the process simply is not prescriptive—the *how* to perform the operations, derive the criteria, or limit the options are not present. The contractor's solution to this state of affairs was a series of rating scales along the various dimensions of both media and skills to be trained. One's confidence in this procedure depends on one's confidence in rating scales. The fact that this process—which included independent judgements of a number of raters—produced a need for a very expensive and not terribly reliable audiovisual device suggests that further refinements to the technique are needed.

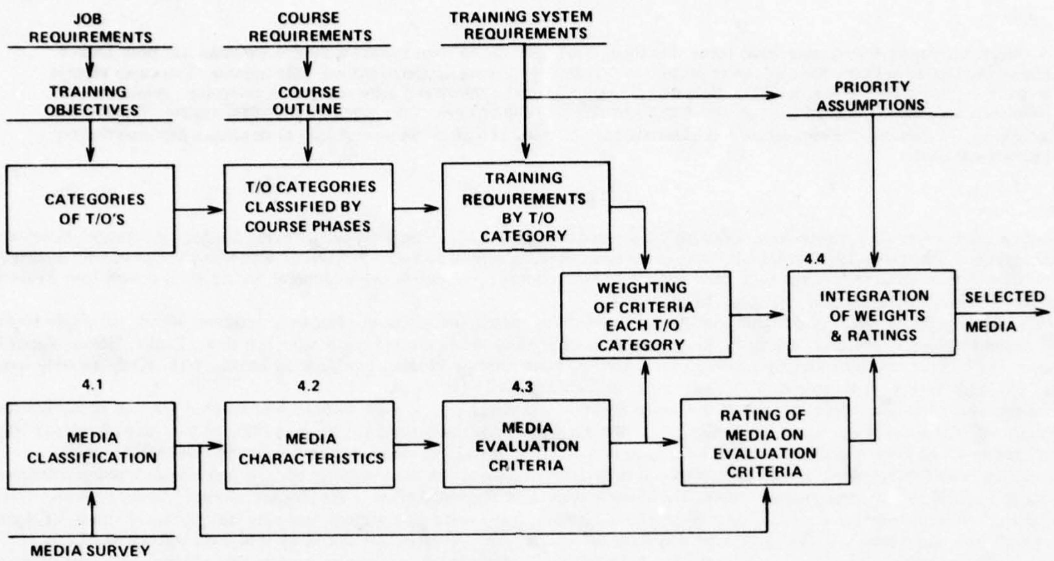


FIGURE 1. MEDIA SELECTION METHOD (From Malone, et al., 1976)

The key to the system is an 8K memory minicomputer with a keyboard and electronic teletype display screen. The input output devices are only for use by the instructors in the present configuration, so this housing for the mini is called the instructor console. The minicomputer drives the Hagan ABC simulator which is a replica of an actual Hagan main console. The only difference in appearance is that the simulator guages are slightly smaller. The student interfaces with the system just as he would on the real thing: by a series of small wheel cranks which he manipulates to initiate changes in the system—one way to obtain information during the troubleshooting process. Also available to the student is a rear projected display of 35mm slides; the computer can display information to the student (e.g., the symptom, feedback on the correctness of a particular choice) or the student can call up information from the system through a set of push tiles on the left hand side of the simulator (e.g., What is the status of some remote component?). The student also presents the computer with his final diagnosis through these push tiles.

There is yet a third fixture to the computer-based system: an eight feet by four feet schematic of an entire automatic boiler control system. It too is driven by the minicomputer and lights come on, go off, or blink slowly or rapidly to indicate various states within the system. Of questionable pedagogical value, it is an impressive device for those of the pinball school of design. In fairness, it does provide a very inexpensive form of animation which can be tied directly to the simulator problems.

A unique instruction capability of the simulator is that if a student misses a problem twice, he is told to call on his instructor and the problem will not continue until the instructor tells it to do so. The system is quasi-adaptive; it cannot handle every idiosyncratic input which students can generate, but it can keep pace with the fastest student and wait patiently for the slowest. The system is built by the Burtek Corporation to the specifications of the Essex Corporation.

PILOT STUDY

The Hagan ABC maintenance course, prior to the implementation of this system, used a stand-up lecture, limited media instructional approach. It required five weeks to complete, all students progressing in unison. Its content, which was not based on a task analysis, was enough different from the new course that comparisons on subject matter tests were precluded. However, time to completion scores could be compared between the old and new courses.

The results are dramatic. Figure 2 shows the distribution of the completion times of the students in the new course as compared to the five week completion time of the old course. Savings in time in excess of 40 percent are seen (Kinkade, Renfro, and Grass, 1978). In a year's time, this would mean a 180 man week savings—almost three and one half man years.

The yearly cost of an E-5 Boiler Technician with ten years in the service (a typical Hagan student) is $14,819 (Martin, Koehler, Mairs, and Hogan, 1977). The savings of three and one half man years is thus nearly $52,000. Since the actual system (hardware, software, and lessonware) cost is right at $250,000 amortizing the cost over ten years results in significant savings, even with the life cycle costs of the system added in. More significant savings can be realized and will be descussed below.

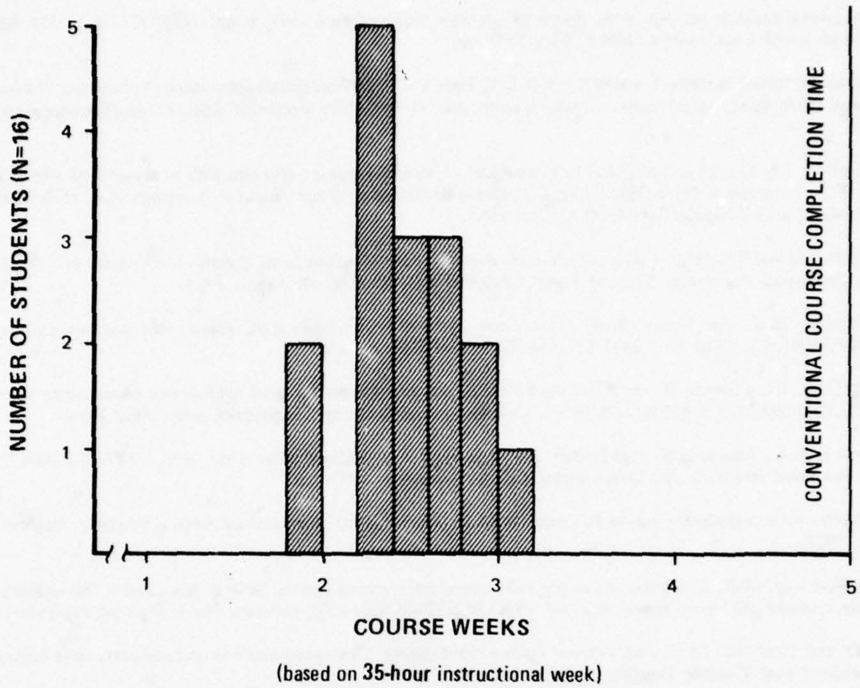


FIGURE 2. STUDENT COURSE COMPLETION TIMES (From Kincaid, et al., 1978)

Course effectiveness was measured in terms of the degree to which stated learning objectives were met. Principal measures of learning objective achievement were obtained from learning module examinations, laboratory exercises, and performance on the simulator. Mean score on the module exams was 95 percent. Laboratory performance was judged adequate by the course instructors; no objective measures were obtained. All students reached criterion on the simulator tests. The mean time to correct diagnosis was a shade under six minutes. Of 23 students in the pilot study, 22 met all objectives in the course.

DISCUSSION

That simulation can be applied in nonelectronics maintenance training courses seems obvious. The minicomputers in this instance were underutilized. Another training system (Woodward Electric Governor maintenance) is being constructed; its simulators will be driven off the same minicomputers, it more than doubles the number of students being trained off the mini's. If time savings comparable to those reported here are realized, real dollar savings could become significant.

The risk in any innovative approach to instruction is that the developed system will be too sophisticated to use once the R&D people and manufacturer representatives have gone home. An integral part of this system is an instructor training package from which the instructors can learn to both operate the system and to alter it software as required. Additionally, one of the specifications for the system was that it use standard Navy parts whenever possible. One possible indication of the benefit of the system is that Atlantic Fegget Hagan maintenance training personnel have requested the system for fiscal year 1980.

It must be stressed that what is reported here is a pilot study. The writer has suggested elsewhere (Brock, 1977c) that measurement of student performance at the exit from a training system is an insufficient index of training effectiveness. Until the students of the new Hagan ABC maintenance training system are measured against students from the conventional course *in the work environment,* the positive results reported herein must be treated as tentative.

REFERENCE

Braby, R., Henry, J.M., Parrish, W.F., Jr., Swope, W.M. *A technique for choosing cost-effectiveness instructional delivery systems* (TAEG Report No. 16). Orlando: Training Analysis and Evaluation Group, April 1975.

Brock, J.F. *Simulation in maintenance training: Now that I've thrown out the bath water, where is the dear baby?* Paper presented to the 1977 American Educational Research Association Annual Meeting. New York, April 4-8, 1977. (a)

Brock, J.F. *Instructional decision making in the design of operator training: An eclectic model* (NPRDC TR 77-31). San Diego: Navy Personnel Research and Development Center, May 1977. (b)

Brock, J.F. Performance testing in instructional systems. In L.T. Pope and D. Meister (Eds.) *Symposium Proceedings: Productivity Enhancement: Personnel Performance Assessment in Navy Systems.* San Diego: Navy Personnel Research and Development Center, 1977. (c)

Brock, J.F. and Delong, J.L. Design and conduct of a mechanical maintenance training program with an annual flow rate of 11,000 trainees: A review. In W.A. King and J. Duva (Eds.) *New Concepts in Maintenance Trainers and Performance Aids* (NAVTRAEQUIPCEN IH-255). Orlando: Naval Training Equipment Center, 1975.

Chief of Naval Education and Training. *A plan for centralized management of instructional systems development with the Naval Education and Training Command. Pensacola: Chief of Naval Education and Training, 15 August 1975.*

Kinkade, R.G., Renfro, R.E., and Grass, W.W. Pilot testing of recently developed ABC Hagan Maintenance Course (NAVTRAEQUIPCEN 76-C-00015C). Orlando: Naval Training Equipment Center, 1978.

Malone, T.B., DeLong, J.L., Farris, R., and Krumm, R.L. *Advanced concepts of naval engineering maintenance training* (Volume I of II) NAVTRAEQUIPCEN N-61339-74-C-0151). Orlando: Naval Training Equipment Center, May 1976.

Martin, J.I., Koehler, E.A., Mairs, L.S., and Hogan, P.F. *Life cycle Navy enlisted billet costs tables—FY77* (NPRDC SR 77-16). San Diego: Navy Personnel Research and Development Center, September 1977.

Merrill, M.D. Psychomotor and memorization behavior. In M.D. Merrill (Ed.) *Instructional design: Readings.* Englewood Cliffs, NJ: Prentice-Hall, 1971.

Miller, G.W. and Rockway, M.R. Maintenance training in advanced instructional system. In W.J. King and J. Duva (Eds.) *New Concepts in Maintenance Trainers and Performance Aids* (NAVTRAEQUIPCEN IH-257). Orlando: Naval Training Equipment Center, 1975.

Montemerlo, M.D. and Tennyson, M.E. *Instructional systems development: Conceptual analysis and comprehensive bibliography* (NTEC IH-257). Orlando: Naval Training Equipment Center, February 1976.

Stewart, W.B. and Kelly, D.M. *Cost and training effectiveness analysis for stand-off target acquisition surveillance system--Final study* (Publication 1979-01-1-1650). Annapolis: Arnic Research Corporation, September 1977.

AVOIDING THE PITFALLS OF ISD

Dr. Melvin D. Montemerlo and Dr. Ward A. Harris
US Army Training Support Center
Individual Training and Evaluation Directorate
Fort Eustis, VA 23604

During the last quarter century, Instructional Systems Development (ISD) has become the pre-eminent tool of instructional technology. The countless thousands of person-hours expended in its application have resulted in the acquisition of a great deal of wisdon. This paper attempts to harness that wisdom by delineating a number of the lessons that have been learned. By taking advantage of this information to avoid the known pitfalls of ISD, present practitioners of the methodology can increase the efficiency and effectiveness of their efforts.

1. Successful application of ISD requires an interdisciplinary team of Subject Matter Experts (SMEs) and skilled instructional technologists. While the need for the former is universally accepted the attempt to do without the latter has been the Waterloo of many ISD attempts. The major thrust of the proceduralized ISD movement has been to reduce reliance on scarce, expensive instructional technologists by developing course design manuals usable by personnel minimally grounded in instructional technology. The development of such manuals, which seek to model the decision making processes of skilled instructional technologists, assumes that: (1) existing theories of learning and instruction are sufficiently refined to permit formulation of course design rules, and (2) instructional technologists agree on those rules. Until both of these assumptions become fact, skilled instructional technologists will be necesary for ISD efforts to be truly successful. The pitfall of failing to recognize this fact can only be avoided by educating managers of training development as to the limitations as well as the advantages of existing ISD methologies.

2. The use of a systems engineering methodology such as ISD requires a great deal of creativity. At most, ISD can be a checklist and a guideline for helping insure that course design errors are detected and excluded before course implementation. An analogy to a similar methodology is helpful. Physics does not tell an engineer how to design a bridge. The engineer must generate and/or collect ideas which can be tested against the laws of Physics to ascertain (as well as possible within the constraint of existing knowledge) the existence of flaws in those ideas. To fill the gaps in existing knowledge, he must use whatever judgment he has acquired through previous experience. The case is similar for the course designer except that the "laws of learning" are much less certain and far less precise. The greater the creativity of the course designer in generating training ideas, the more likely the resulting course will be successful. Instructional technology can be used to critically examine those ideas but rarely to generate them. Compare, for example, the highly successful American Airlines pilot training program or the televised Sesame Street program to courses which have been generated by SMEs using an ISD manual. The best evidence that course design personnel possess the needed creativity is not the possession of advanced academic degrees or of ISD workshop graduation certificates, but "track records" of having completed similar projects successfully.

3. Adaptation is the rule rather than the exception in the use of instructional design models. For example, ISD manuals generally provide only one or two procedures for task analysis. However, task analysis is a means toward an end and not an end in itself. The end may be a new training program, revision of an existing program or redesign of the existing job. In each situation, the end to be accomplished must be examined, the information needed to reach that end must be identified, and a plan to collect the information must be developed. Once implemented, the plan must be monitored and, if necessary, modified. Since no algorithms exist for adapting, monitoring or modifying them, these processes constitute an art rather than a science. As such, they are best learned in apprenticeship programs under the guidance of those with demonstrated competence. Implementation of apprenticeship programs for novice instructional technologists will insure the availability of sufficient numbes of "masters" in the future.

4. The use of an unchanging tool such as an ISD manual in a field which changes as rapidly as instructional technology causes a "wisdom lag." Unless course designers are professionals who constantly strive to keep up with the work of their colleagues and with the literature on training technology and educational psychology, they are not making use of the best information available. The less experienced the ISDer, the more dependent he is on the manual, and the less aware he will be of its limitations. Organizations wishing to minimize this wisdom gap should select and develop a cadre of professional instructional technologists and provide for their continued professional develomment.

5. The required use of a specific ISD manual results in a centralizing bias in the courses which are developed. This would be desirable if the bias was toward a good norm. However, since no ISD manual has ever been tested versus any other manual or versus no manual at all, there is no way to determine whether the centralizing bias is for better or for worse. Organizations wishing to achieve an uplifting rather than a centralizing bias should emphasize the products rather than the processes of course design. This can be done by specifying desired course parameters and letting course designers select or design methods for achieving them.

6. ISD can become a ritualistic, that is, a rule-following rather than a problem-solving process. The problem is exacerbated if the ISD team perceives they are more likely to be punished for failure to follow the rules than to be rewarded for doing whatever necessary to solve the problem. The lower the rank/status of the ISDers, the more likely this phenomenon will occur. Straying from the procedure requires the knowledge of alternatives, the confidence to implement them, and the perception that that is what is wanted. To inhibit the ritualization of course design procedures, an organization must make it clear that results, not ritual, are desired.

7. Although ISD manuals do not mention it, course design is a highly political enterprise. Any process that innovates will provoke anxieties in those with vested interests. To be successful, a course design effort must have the complete support of the organization's top manager, and his subordinates must be aware of it. Even then, the course design team must walk a political tightrope, using sufficient force to get the job done and sufficient flexibility to avoid alienating the instructors and middle-managers who can easily subvert the new course after the design team leaves. During task analysis, the team must be able to discern when the "wool is being pulled over their eyes" and to rectify the situation fully yet amicably. The necessary political savvy and judgment comes with experience in some cases, and in other cases may never come at all. A charismatic personality is not without value to a training technologist.

8. No institution can be expected to perform an incisive, unbiased analysis of their own needs and policies. While it is not within the nature of organizations to "hangout their dirty wash," it is absolutely necessary for problems to be aired if a systems analysis is to be meaningful. To avert this problem, the design team leader should be an "outsider" to the organization because he would be more likely to spot institutional biases, and because he would be less likely to be hurt for surfacing them.

9. The scientific, objective aura surrounding ISD renders it highly suitable for covering up the implementation of self-serving administrative decisions. This potential abuse of all systems-analytic methodologies has long been known. In ancient times, leaders found their proclamations had more effect if it was announced that they came from an oracle. In modern times, the same effect is achieved if the proclamation is announced to be the result of a systems analysis. The best guarantee this is not happening is the reputation of the ISD team leader.

10. In large complex established organizations, it is difficult to implement a meaningful ISD because administrative control is spread across a network of interwoven sub-organizations with separate responsibilities. An ISD team may find that changes should be made to doctrine, manuals, equipment and organizational structure, and then find that the sub-organization to which they report has no control over these factors. These problems can be minimized if the ISD team includes persons well-versed in the workings of the organization.

11. The funding cycles of large organizations are often out of phase with the timing of the ISD. In the military, the funds for large ISD programs must be programmed long before the study is begun and even longer before its recommendations are known. To compensate for this, the cost of expensive simulators and other media can be built into the programmed funds. Since it is not within the nature of such organizations to return appropriated funds, the results of the study can become a foregone conclusion.

12. Since humans rarely act contrary to their own best interests, ISD does not always produce efficient training programs. In industry, where training expenses reduce profits, the contingencies of reinforcement favor the reduction bottom line and where two powerful measures of an officer's importance are the size of his budget and the number of his subordinates, the contingencies of reinforcement are reversed.

13. Overemphasis on terminology can lead to naive and to deliberate misuses of the technology. A shallow understanding of task analysis, for example, can lead well-meaning personnel to gather much more data than the ISD team can use. The problem of deliberate misuse is exemplified by the ISDer who uses the buzz-words for purposes of obfuscation (e.g., to hide the preservation of the status quo or the implementation of predetermined policies). These problems can only be ameliorated through longer and more intensive training of ISDers and managers at all levels.

14. The term "task" is often used in the ISD process, yet it has no universally accepted referrent in reality; that is, there is no optimally-sized chunk of behavior which always facilitates course design. The term "task" is merely a tool for partitioning global or molar behaviors into conveniently sized ones. The size that is convenient may vary greatly within a course; between courses the variation knows no bounds. Long hard debates on the size of a task are part of the initiation rites for training technologists. When one no longer feels the need to enter such arguments, that part of the initiation is over. By including such initiation rites in the training of ISDers, the number of wasted hours in actual course design projects can be greatly reduced.

15. Very few skills are validly analyzable down to a series of stimuli and responses. While simple procedural tasks can be so described; others cannot. Those that are difficult to analyze include "soft skills" such as leadership, complex skills such as troubleshooting, intellectual skills such as training program design, and psychomotor skills such as buttoning a sweater. Stimulus-response flowcharts, while useful in describing the relatively simple and choppy behaviors of a novice, may be highly misleading in describing expert performance of the same skill. Novices in instructional technology finish the second portion of their initiation when they realize that it is neither possible nor useful to break all tasks down into a stimulus-response chain of non-subdividable components. Bringing about this realization during the training of ISDers is much more efficient than allowing it to occur on-the-job via the discovery method.

16. It is neither possible nor desirable to break down the over-all objectives of a training program to an extremely large number of atomistic specific behavioral objectives (SBO). SBOs should be used for guiding course design and should not become overly voluminous. However, the stopping point is judgmental. The urge to develop SBOs ad infinitum is best brought under control during the formal training of ISDers.

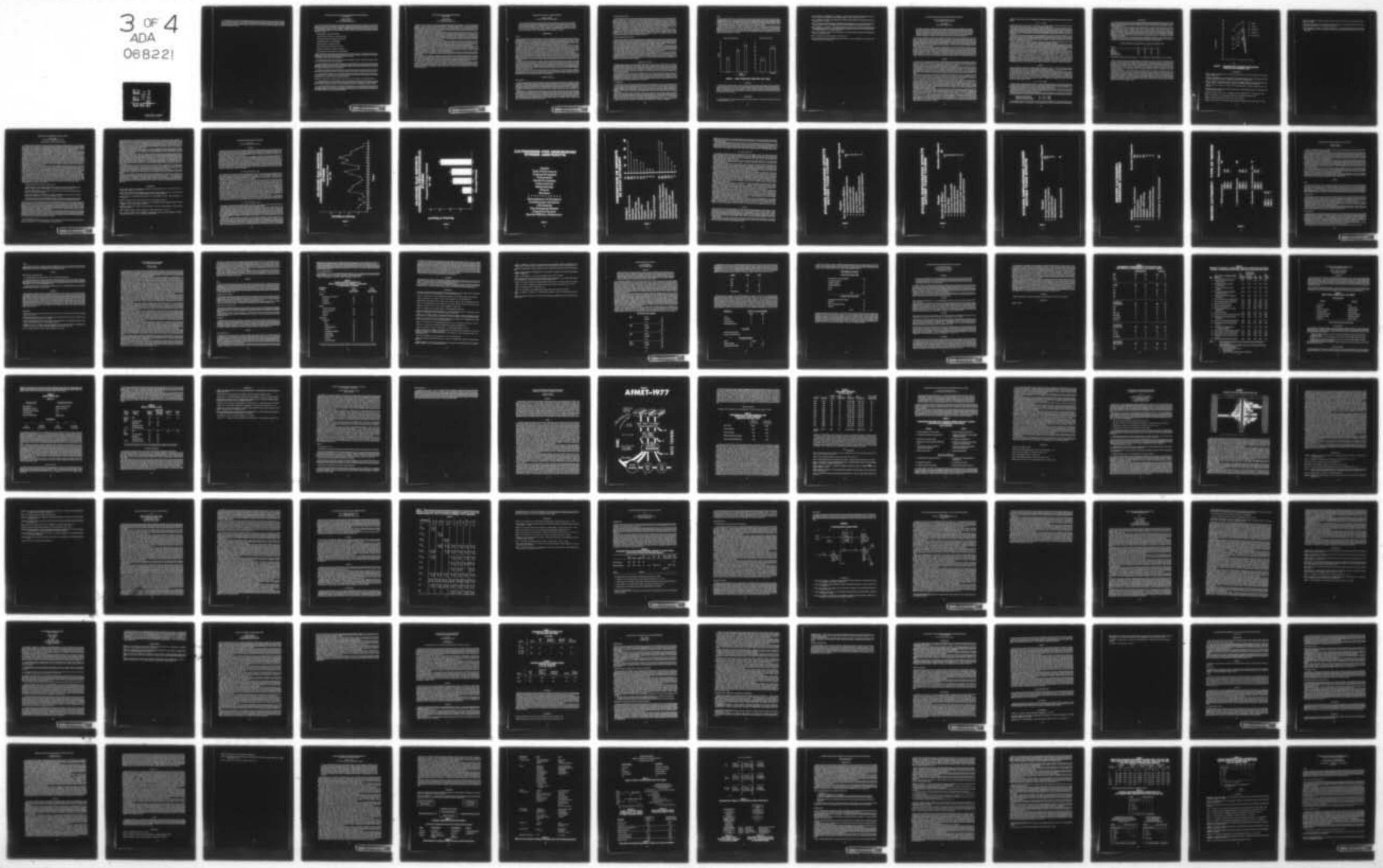
17. The definition of training objectives is not a straightforward procedure. It is likely to be the most heart-rending portion of course design. The simplest, cheapest, and most circular method of determining what should be taught is to ascertain from existing course documentation what is taught. Actually going to the field to determine what skills shoulds be taught is expensive, time-consuming and usually frustrating because the SMEs can't agree. The answer lies in accepting the fact that disagreements will occur and the fact that the resulting statement of objectives will be, at best, a compromise. The most important skills to be learned for any job are impossible to define precisely. They include the ability to generalize from what has been learned to whatever may arise on the job. Psychologists refer to this as "learning how to learn" or "generalized transfer of learning." An excellent method of learning about the problems inherent in the definition of training objectives is to review the failures, the tribulations and the false starts as well as the successes of others.

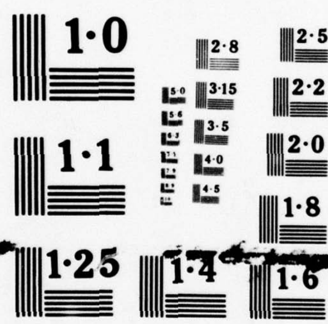
18. The implication of reductionistic and deterministic theories of instruction, such as ISD, is that instruction is simply the injection of a pre-defined sequence of information into the mind of a passive learner. To an even greater degree than ISD, this was the implication of the Programmed Instruction (PI) paradigm that died of its own sheer weight. The work of Bruner, Bloom, Montesorri, Piaget, Whitehead, and others had shown this position to be untenable long before it had been adopted by either PI or ISD. If there is a single fact that the changing theories of education through centuries has produced, it is that human beings are so adaptive that they will learn regardless of the teaching method used (if they want to). The natural state of the human being is that of learning. Indeed, it is more difficult to keep humans from learning than to force them to. The real work of the instructional developer is to facilitate the learners job by bringing together and, to some extent, organizing information, and by giving him a chance to practice and receive feedback. The learner will do the rest if he desires. To think otherwise is the ultimate conceit for the instructional technologist. But then again, everyone likes to feel needed.

19. There is not "one best way" to teach a skill, much less an entire course. The purpose of front-end analysis is not to determine the best way, but to insure that as many mistakes are avoided as possible. If any two sets of course developers independently undertook the development of a course, the resulting courses would be highly dissimilar even if they were using the same ISD manual. However, since available tools for measuring the goodness of a course are not very powerful, it is doubtful the measures would be implemented because of their cost. Furthermore, because organizations don't like to admit mistakes, it is unlikely the measures would be used to search for them. Anyone who doubts this should review the literature on training device evaluations. They will have to look long and hard to find indications of error in design. Although "quality control" systems are often mentioned in ISD manuals, those that have been implemented are designed only to make minor changes. For the most part, training program design is a one-shot affair. Unfortunately, for better or for worse, many ISDs are doomed to success. An excellent method of learning that there is not a single best way to teach a skill is to have a group of novice ISDers independently design, in a classroom environment, a course to teach the same objective. Examining each others resulting courses and considering how difficult and expensive it would be to determine whose course is best, are powerful tools for developing a realistic perspective on the advantages and limitations of modern training technology.

AD-A068 221 AIR FORCE ACADEMY COLO F/G 5/10
PROCEEDINGS OF THE ANNUAL SYMPOSIUM ON PSYCHOLOGY IN THE DEPART--ETC(U)
APR 78 J M BERMUDEZ, W H CLOVER, L J DAHLE
UNCLASSIFIED USAFA-TR-76-12 NL

3 OF 4
ADA
068221




1·0
2·8
2·5
3·15
2·2
3·5
2·0
4·0
1·1
4·5
1·8
1·25
1·4
1·6

20. ISD requires more time, money and personnel than traditional course design methods. It should not be undertaken unless there is reason to believe that the increased cost of the design will be more than offset by a savings in the resulting course. This is more likely to be the case with courses which have a high per-student cost, a high student flow, and a long stable course life. It should be remembered that since the beginning of humanity, people have passed their skills onto others. The most that instructional technology can do is to help SMEs develop more efficient courses.

AIR FORCE SKILLS MAINTENANCE AND REACQUISITION TRAINING PROGRAM (Project SMART)

Edward E. Eddowes
Flying Training Division
Air Force Human Resources Laboratory

The Air Force (USAF/XOOD) had identified a priority research requirement for the development and validation of an Air Force skills maintenance and reacquisition training program (Project SMART). This research will develop and validate comprehensive, objective, quantitative procedures which will permit the efficient management of individualized flying training and provide mission readiness at minimum cost. Project SMART was initiated and planned by AFHRL in anticipation of future Air Force flying training requirements where energy and aircraft costs preclude application of current training practices. Project SMART will provide the following products of substantial benefit to the Air Force:

- Recommendations for improved flying training
- Analysis of training cost trade-offs in maintaining flying skills
- Test of alternative policies for management of the rated force
- Determination of factors in pilot skill degradation
- Operational definition of pilot mission readiness
- Objective measurement of individual and unit readiness
- Development and evaluation of innovative training methods
- Improved procedures for classiflying pilots for flying assignments

Because the prospects for reduced flying time authorizations are more critical for the Strategic Air Command (SAC) and Tactical Air Command (TAC), and because the missions of these commands require a wide variety of flying skills, Project SMART has been designed to focus research on the critical flying skills required by SAC and TAC missions.

Project SMART consists of these six program elements:

(1) Program Development and Planning: This effort initiated Project SMART and will define, document and coordinate the effort for review and approval by higher Air Force headquarters.

(2) Preliminary Evaluation of Project SMART Methodology: Critical Project SMART research elements will be selected and tested to evaluate the concepts and methods of the program. The preliminary evaluation will provide timely feedback information in subsequent research phases. In addition, the evaluation will provide early indication of the kinds of results which may be expected and how the results may be applied in optimized future Air Force flying training.

(3) Identification and Definition of Critical Flying Skills: The critical flying skills required for operational proficiency in tactical and strategic missions will be identified and defined. The results of this analytical effort will furnish an operational definition of mission readiness.

(4) Develop Techniques for Measurement of Critical Skills: Using the flying skill definitions generated during the analytic study as a guide, techniques will be developed to measure the flying skills required for mission readiness. The skill measures will be objective and quantitative wherever possible and must be operationally meaningful.

Development of a suitable array of skill measurements is critical to the success of the maintenance of the flying skills program.

(5) Measure Skill Retention: With skill measurements available, a series of studies will be accomplished in which the loss or decrement of mission readiness flying skills is recorded as a function of the length of time during which no flying occurs. The output of this program element will describe the extent of the problem of maintaining flying skills by defining the skill decrement that must be avoided.

(6) Dvelop and Evaluate Skill Maintenance and Reacquisition Training Programs: Once the skill degradation that results from periods of no flying is determined, cost effective means for maintaining the required skill levels or for retraining to eliminate the skill loss must be developed and evaluated. The more promising of the maintenance and retraining programs will be tested further under operational conditions to validate the effectiveness of the skill maintenance program.

PILOT PERFORMANCE MEASUREMENT SYSTEMS
(Project 2359)

Wayne L. Waag
Flying Training Division
Air Force Human Resources Laboratory

Within the Air Force, there has occurred an increased emphasis on flight simulation training primarily as a result of spiraling aircraft operting costs. In order to document the effectiveness and efficiency of the simulator training syllabus, it is desirable to quantitatively assess aircrew proficiency both in the simulator and the aircraft. Once a measurement capability exists, it is possible to precisely determine the effects of changes in the syllabus of instruction, and therefore, optimize the flight simulation and aircraft training programs. Such a system would also allow the precise definition of proficiency requirements. In this manner, the flying training manager could more readily control the quality of graduates from his training programs.

The development of objective performance assessment capabilities for operational flight training systems will most likely proceed in two phases. The first phase will focus on the implementation of measurement systems within the flight simulation environment while the second will focus on the aircraft. There are definite advantages in accomplishing the initial development efforts in the simulator. Since the simulator environment can be precisely controlled, the development and validation of assessment algorithms can readily be accomplished. Likewise, the flight simulator provides access to all control input and aircraft state parameters whereas the number of parameters sampled in the aircraft environment is usually limited. Selecting the most critical parameters based on simulation data should save time and reduce costs in the development of an airborne measurement system.

The Air Force Human Resources Laboratory has recently initiated a 5-year program for the development and installation of objective performance assessment systems in selected flight simulators. The program has three major thrusts. First, an automated performance measurement system will be developed and implemented for selected simulator systems. Since continuous performance monitoring is impractical, the goal is the development of one or more automated mission profiles which could be used to assess aircrew proficiency.

A second major thrust will be an evaluation of the utility of automated performance measurement in operational training. First, the correspondence between measured performance in the simulator and aircraft will be determined. The prediction of performance in the aircraft from performance in the simulator is viewed to be a major potential benefit of the program. Second, the measurement system would be evaluated according to the usefulness of the feedback provided to the student and instructor. And third, it would be evaluated according to the usefulness of the information provided to the syllabus designers.

A third major thrust will be the development of a set of functional specifications for the inclusion of a performance measurement capability for future generation devices. It is expected that requirements would also be generated for the development of retrofit capabilities for existing simulation systems.

Several criteria have been established for the selection of simulation systems for which a measurement capability would be developed. First, it is desirable that the systems represent a broad range of aircraft types and missions. Given the limited resources, it is important that the results be as generalizable as possible. Furthermore, the simulation systems, per se, should represent different levels of "state-of-the-art". At least one system should require a retrofit, while another should be able to support the implementation of the measurement system using existing hardware. Where possible, it seemed that the selected systems should build upon existing technology.

The focus of the 5-year program just described is the development of objective proficiency assessment systems for flight simulation. Nevertheless, pilots are trained to fly aircraft, and it is this environment for which objective proficiency assessment techniques are ultimately needed. It is anticipated that the long range performance measurement efforts will be directed toward providing this capability in the aircraft. The successful development and implementation of objective assessment in the simulation environment should pave the way for airborne performance measurement.

DESIGN AND EVALUATION: A SYSTEMS APPRAOCH

Dennis B. Beringer
University of Illinois at Urbana-Champaign

A coherent and logical approach to design and evaluation of complex man-machine systems should include a system analysis, synthesis, and functional allocations determined during analysis are combined with relevant human factors and engineering data to synthesize system configurations. A candidate method for evaluation, the Central-Composite Design (CCD), is briefly reviewed. Both systems approach and CCD evaluation are applied to a 4-D navigation control display system, showing how the CCD can be used to evaluate system operator performance systematically and economically across a wide range of task and environmental variables. Preliminary findings and their implications are discussed.

INTRODUCTION

Problem

During earlier periods of human evolution, the systems created were of limited complexity. The implementation of design modifications typically proceeded in a trial and error fashion that was, in most instances, sufficient for changes of limited scope. However, as systems have attempted operations of ever increasing complexity, the process of introducing improvements has itself become a problem. Contemporary systems have reached a level of complexity that makes trial and error solutions both uneconomical and impractical. What is needed is an efficient, economical, and credible strategy for facilitating the evolution of complex operational systems.

The evolution of many systems can be facilitated by the systematic application of simulation in an iterative process of analysis, synthesis and evaluation. The National Airspace System (NAS) is one for which such an approach may offer the only practical solution. The effective functioning of this system is threatened by increasing demands on its capacity and environmental restrictions on its range of operation, particularly in congested terminal areas. Increasing automation is being called upon to expedite safe and efficient traffic flow in the face of these demands. Both the functions and the modes of operation of the major subsystems that comprise the system, including airborne flight control (pilot and aircraft) and ground flight control (air traffic controller and computer), will be affected dramatically.

Not only are the major subsystems in use facing reevaluation, but also the tasks each performs. At present aircraft flight paths are assigned in three dimensions within contraints imposed by complex networks of airways and terminal area procedures. Departures, enroute progress, and arrivals are adjusted by traffic controllers and pilots through various "holding" procedures and relatively gross groundspeed limitations in accordance with a body of separation criteria intended to assure freedom from traffic conflicts in areas of positive control and reasonable safeguards in other areas. Such procedures achieve safety through operationally undesirable and uneconomical restrictions in traffic flow.

The adoption of explicit, time-phased (4-D) "metering and spacing" procedures that depend upon precise compliance by all aircraft will allow the maximum reduction in separation criteria without air traffic conflicts while providing emergency procedures to accommodate rare but inevitable departures from flight plans and clearances. The implementation of such procedures, however, depends upon a major transfer of functions from controllers to computers and the consequent re-definition of the controller's job to include certain new functions not required in current operations. To accommodate such functional changes will require new means of processing and displaying information, to both controllers and pilots, and new means of communication between and among men and machines (Carel, McGrath, Hershberger, and Herman, 1974; Fenwich, 1970).

The focus of special interest is the interface between the pilot and the navigation computer. It has become necessary to reshape the means by which flight crews enter and retrieve data using advanced electronic navigation display systems. The approach being used in this investigation involves a systematic analysis of the requisite tasks, a coherent system synthesis, and an efficient and economical evaluation strategy which relates task and environmental variables to system performance.

A SYSTEMS APPROACH

Major Components

A coherent and logical approach to design and evaluation of complex man-machine systems should include a system analysis, synthesis, and evaluation. The system analysis should include a statement of system goals, identification of necessary subgoals, identification of information necessary for system control, and identification of functional operations required for control. The system synthesis should follow, beginning with a man-machine functional,allocation. Applicable design principles and restrictions should then be identified, required information organized for integrated presentation, and orders of control authority of man defined. After these steps have been completed one or more candidate system configurations should be developed which will allow comparative evaluations of various costs and trade-offs to be made.

A meaningful system evaluation should use tasks which are representative of the system's expected range of operation and a methodology which allows assessment of system performance across wide ranges of a number of variables. The great majority of evaluations of complex system performance have isolated certain elements of a particular task for examiniation. These are often called reduction experiments. In many instances this type of approach may not allow environmental variables as they ineract across their full ranges. It is also often the case that large portions of the experimental variance are not accounted for (Simon, 1976). Although systematic measurement of the effects of previously excluded factors might increase the amount of variance accounted for, the addition of factors to a factorial design could cause the design to becom unwiedly in terms of is size.

A Candidate Experimental Design

A reasonable estimate of the functional relationship between task variables and human performance may, however, be obtained by approximating the "response surface," an approach known as *Response Surface Methodology* or RSM (Box and Wilson, 1951; Myers, 1971). RSM frequently involves the use of second-order central-composite designs (CCD). The CCD dictates the data points are designated by adding strategically selected factor combinations to a basic two-level factorial design. These additional points, collectively referred to as the *star* or *axial* portion of the CCD, are selected with regard to two desirable properties; rotatability and orthogonality (Clark and Williges, 1973). The design is rotatable if the variance of the predicted response is the same for all data points equidistant from the center. This is desirable when the orientation of the response surface relative to the orthogonal factor axes is unknown. Orthogonality allows data points to be separated in orthogonal blocks, any differences in performance among blocks being indepenent of main effects.

The CCD may be analyzed in several ways depending upon the allocation of variables to within-subject and between-subjects categories. The usual evaluation would involve a multitude regression analysis followed by an analysis of variance of the regression coefficients (Clark and Williges, 1973). If some within-subject factors are used, means across subjects may be submitted to a regression analysis or individual subject scores may by analyzed by regression techniques followed by an analysis of variance. It is also possible to conduct an analysis of mixed-factor CCD design and analysis are presented by Dexon.(1976).

Environmental Manipulation Using the CCD

Systems and devices, the design of which are influenced by human factors engineering efforts, must somehow be examined in the context of the real-world environment in which they will find application. A systematic and economic strategy for this evaluation uses the CCD to structure combinations of environmental variables that are expected to influence system performance, including variable ranges expected to occur in the operational environment. Although the CCD has occupied a *central* position in some previous human factors investigations, it may take an *ancillary* role for the present purpose as illustrated by Eisele, Williges, and Roscoe (1976).

Eisele *et al* examined pilot judgments of lateral and vertical deviations from a four-degree landing approach as influenced by five display factors. Four contact analog display elements (runway outline, runway touchdown zone, runway centerline, and ground-plane texture grid) and one synthetic display element (a modified "highway in the sky") were manipulated in the context of a two-level factorial design. Each of the resulting 32 display conditions was examined across 27 different flight position and attitude viewpoints generated by the manipulation of five variables (range, vertical deviation, lateral deviation, pitch, and bank) within the context of a CCD.

The value of including these positional and attitudinal variables in a CCD manipulation is twofold. First, even if these variables are not examined in terms of task performance influences, a gain in generality of results has accrued due to the *systematic* sampling of environmental variables that might affect the discrimination task. The effects of the display variables need not be interpreted within the context of a single environmental set. Second, the data related to the CCD may be submitted to regression analyses to examine the quantitative relationships between environmental variables and task performance. In this way some meaningful correspondence may be developed between expected task or real-world states and task performance.

APPLICATION: 4-D RNAV

The suggested system analysis and synthesis were applied to a pilot/computer interface for 4-D area navigation (RNAV). The application of human factors design principles and the reconfiguration of task structure for optimum allocation of functions produced a new system configuration embodying several significant advances. Display graphics, display integration, and direct displays-operator interaction methods were used to produce the integrated touch-access control/display unit, combining a horizontal situation display (HSD), vertical situation display (VSD), and direct manual accession device (touch panel). Most significantly, the inclusion of direct manual accession eliminated the need for peripheral intermediate symbolic devices (keyboards) to enter or retrieve data crucial to system operation, allowing procedures to be more direct and intuitive.

This system configuration and two others have been simulated on the PLATO IV computer system at the University of Illinois. The other configurations use a simulated conventional entry keyboard and either a static or dynamic map display. Conventional error indicators and flight instrumentation have also been generated on the system and are used to simulate the aircraft. These three configurations represent increasing degrees of computer aiding or automation of the data processing tasks. The touch system represents the highest degree of aiding, allowing inflight route modifications to be made by the direct manual modification of the map display. The great advantage of this system is that graphic solutions to navigation problems may be entered directly.

CCD Implementation

It is not sufficient to select any single navigation task and compare system A with system B for that task. A far better approach would be to retain the system variable as a separate between-subject factor and vary the difficulty of complex navigation tasks (4-D RNAV: x, y location; altitude; time) by manipulating within-subject task variables using a CCD. The task and environmental variables selceted (command groundspeed, vertical gradient, degrees of turn, drift across course, drift along course) should produce, independently or in combination, differential degrees of task difficulty. The dependent measures related to task performance include both flight parameters (horizontal, vertical, and time errors) and system entry parameters (blunders, entry times, ATC procedural compliances, and side-task response rates).

In the five-factor case the factorial portion of the design is a 1/2 replicate of a 2^5 factorial. The values of alpha for rotability and orthogonal blocking are both 2.000 (Clark and Williges, 1973). This selection allows both properties to be retained and permits the possible inclusion of quasi-continuous variables that assume integer values should it become desirable at any point. The complete CCD contains 16 factorial data points, 10 axial points, and 6 center points. The factorial 1/2 replicate contains coded factor levels of +1 and -1, the axial portion +2 and -2, and the center points zero. Thus the response surface can be approximated for 5 factors at 5 levels using a grand total of 32 data points (27 unique factor combinations). A full factorial design would contain 3125 unique factor combinations.

Analysis

The collection of performance data for each navigation system configuration, using independent groups of 8 pilots, allows the development of regression equations for each, relating system performance to task and equipment variables. Areas of relative strength and weakness and differences between systems in the magnitude of factor influences are expressed explicitly in quantitative terms. The analysis includes a multiple regression analysis for each dependent variable based upon the five independent variables and subsequent analysis of variance of the regression equations. The variances of secondary-task response rates should be indicative of differences in residual attention resulting from differential primary task difficulty in operating the conventional and the more advanced systems.

Presently only preliminary procedures tests (Figure 1) using input device are available. The results indicate that the touch panel system is slightly faster than the keyboard system for recalling a single stored waypoint (1.7 and 2.7 seconds respectively). The touch panel proved highly superior to the keyboard, however, *in the entry of data to create a new waypoint* (15.7 seconds and 30.0 seconds respectively). (Dotted lines represent a standard deviation about the mean.) This difference undoubtedly reflects the reduced number of motor responses required by the touch system achieved through procedural and operational simplification and suggests that marked differences may be found during the simulated flight performance evaluation.

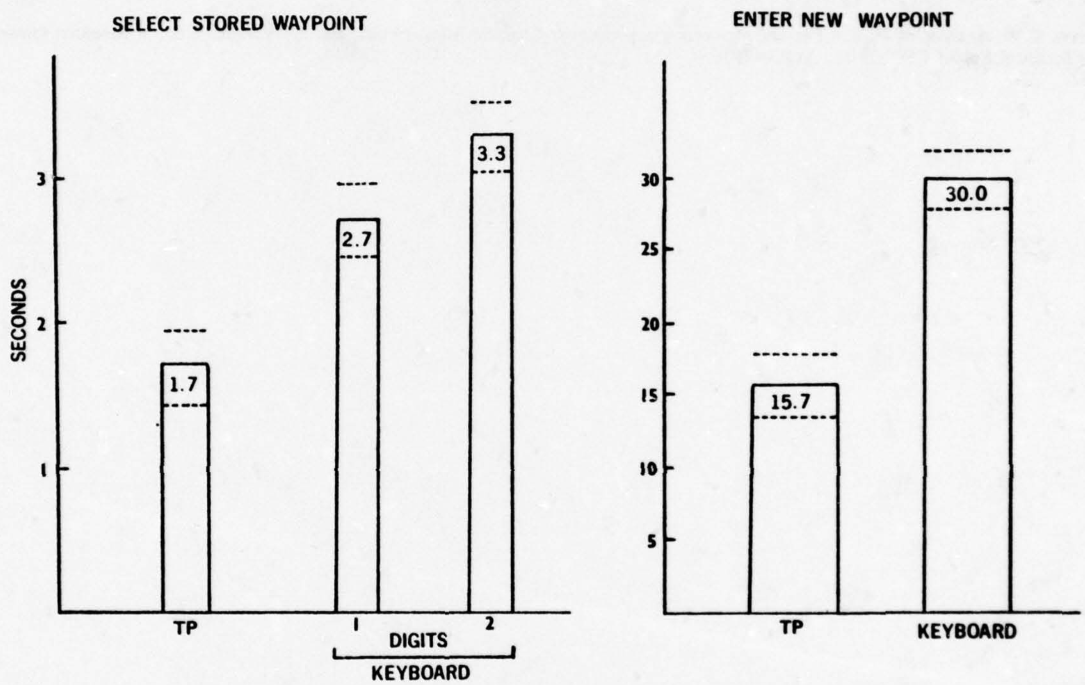


FIGURE 1. MEAN COMPLETION TIMES FOR TWO TASKS

SUMMARY

It is apparent that the increasing emphasis on energy conservation and budgetary frugality will bring greater numbers of experimenters to seek the most efficient means for screening a host of equipment design, training, and procedural variables that influence human performance. Equally important will be the economic assessment of the impact of task and environmental variables upon manned system performance. The method outlined here is one possible solution, not only to the problems presented by an evolving National Airspace, but generally to the problem of efficient experimental investigation, providing a systematic and economic strategy by which the design or evolution of any system may be greatly enhanced.

REFERENCES

Box, G.E.P. and Wilson, K.B. On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society*, Series B (Methodological), 1951, *13*, 1-45.

Carel, W.L., McGrath, J.J., Hershberger, M.L., and Herman, J.A. Design criteria for airborne map displays. Hughes Aircraft Co.: Joint Army-Navy Aircraft Instrumentation Research Program, Report No. 731101, Culver City, CA., 1974.

Clark, C. and Williges, R.C. Response surface methodology central-composite design modifications for human performance research. *Human Factors*, 1973 *15*, 295-310

Dixon, C.C. Mixed-factors central-composite designs: A theoretical and empirical comparison. Savoy, IL.: University of Illinios at Urbana-Champaign, Institute of Aviation, Aviation Research Laboratory, Technical Report ARL-76-13 AFOSR-72-5, June 1972.

Eisele, J.E., Williges, R.C., and Roscoe, S.N. The isolation of minimum sets of visual image cues sufficient for spatial orientation during aircraft landing approaches. Savoy, IL.: University of Illinois at Urbana-Champaign, Institute of Aviation, Aviation Research Laboratory, Technical Report ARL-76-16 ONR-76-3, November 1976.

Fenwich, C.A. The pilot interface in area navigation. Cedar Rapids, Iowa: Collins Radio Company, paper prepared for the Human Factors Society 14th Annual Meeting, 1970.

Myers, R.H. *Response surface methodology*. Boston: Allyn and Bacon, 1971.

Simon, C.W. Analysis of Human Factors engineering experiments: Characteristics, results, and applications. Canyon Research Group: Technical Report CWS-02-76, August 1976.

MULTIFUNCTION SWITCHING AND FLIGHT CONTROL WORKLOAD

Billy M. Crawford and William H. Pearson
Aerospace Medical Research Laboratory

Mark Hoffman
Systems Research Laboratory, Inc.

Four subjects were tested in a cockpit simulator using a secondary task to measure reserve information processing capacity under two levels of flight control and four levels of multifunction switching. Results suggest that flight control impacts both input-output and central processing stages whereas mere anticipation of switching tasks affects input-output only. Difficult flight control reduced the effective information processing reserve by 54 percent on the average. The corresponding losses attributable to anticipation of multifunction switching were 20 and 31 percent for simple and complex tasks respectively. The study has implications for design of effective digital processing aids and mental workload measurement.

The evolution of compact digital computers has made possible the development of digital avionics information systems. Such systems promise a number of advantages to both aircraft designers and users. For example, when interfaced with multipurpose cathode ray tube displays and multifunction switches, digital computation and stroage capabilities can be used to reduce the number of dedicated instruments competing for cockpit panel area. Information which is not required by the pilot on a continuous or frequent basis can be stored and presented on demand either automatically, as related programmed mission events transpire, or in response to manual control actions (Zipoy, Premselaar, Gargett, Belyea and Hall, 1970). And with reduced demands for panel space, it will be easier to locate the multipurpose controls and displays in prime reach and viewing areas.

However, experienced pilots are troubled by the prospect of possible added activity—both mental and physical—required to gain access to information which is normally on dedicated instruments. Should the demand for such activities occur during peak operator workload, the impact on mission success might not be offset by the increased calculating power, speed, or accuracy afforded by the digital processor. One of the purposes of this study was to evaluate the impact of multipurpose control display tasks on the pilot's reserve capacity. Of particular interest was the question as to whether or not the maintenance of knowledge of procedures associated with multifunction keyboard operation reduced the operator's reserve capacity for making choices or decisions such as might be required to handle contingency situations during a mission. Another purpose of this study was to investigate the compatibility of keyboard operations with continuous flight control tasks.

METHOD

Apparatus and Tasks. A computer-based simulator was used to present and score the task situations investigated (Brandt and Wartluft, 1975). Of the three different tasks involved, two, flight control and communications IFF switching functions, represented actual tasks in aircraft systems. The third was an information processing task which served as a test to measure cognitive reserve capacity under various primary task conditions. All three tasks were implemented within a fixed-base cockpit simulator.

The front panel of the cockpit was equipped with three CRT-type displays. The center display was used to present information concerning basic flight parameters in a moving tape format.The The cockpit also contained a throttle with afterburner switch (left side panel) and a center-mounted joystick control which were used, in combination with the displayed flight information, to "fly" various maneuvers. Printed computer outputs of simulator performance data included both mean absolute and root mean square error relative to specified control values based on "fly to" instructions for altitude, heading, bank angle, pitch, indicated airspeed, vertical velocity, angle-of-attack, and g-load.

Between the front instrument panel and left side panel was a multifunction keyboard (MFK). This MFK, in combination with the CRT on the upper left of the front panel and a numerical entry keyboard, which was also located on the instrument panel (lower left), was used to simulate a multifunction interface with digital avionics subsystems. Subsystems, functions and states were displayed on the CRT to complement the feedback afforded by back-projected legends on the MFK pushbutton faces.

The reserve capacity test used was a variation of the Sternberg choice reaction task (Sternberg, 1966,1969). This particular task was selected on the basis of its use by Briggs, Fisher, Breenberg, Lyons, Peters and Shinar (1971) and Briggs, Peters and Fisher (1972) in studying divided-attention effects, i.e., the tendency for subjects who perform two tasks simultaneously to be less proficient on one (if not both) than they are when performing the tasks separately. Briggs attempted to localize the divided attention effect within a four-stage model of human information processing: (1) encoding of stimulus information; (2) central processing to identify and classify the stimulus; (3) response decoding; and (4) response execution. The Sternberg task allows the researcher to vary central processing demands while holding input and output requirements constant.

The Sternberg task procedure used in this study was as follows: At the start of an experimental session, the experimenter read to the subject a set of 1, 2, 4 or 6 letters of the alphabet. The subject was asked to retain the set in memory during the succeeding block of trials. The four sets used were as follows: A, AH, AHJQ,and AHJQSX. Such sets are referred to as "positive sets." During the block of trials the subject series of test stimuli or "probes" to which he was to make one of two responses: (1) "yes," the test stimulus matches the positive set, or (2) "no," it does not match, and, hence, is a member of a negative set. The negative set included the 9 letters, B, C, E, F, G, I, L, R and Y. Negative and positive stimuli occurred with equal probability (.5). Letters within the two sets also occurred with equal likelihood. The average inter-stimulus interval was 5.5 seconds and ranged from 3 to 7 seconds. "Yes" was indicated by the subject's pushing forward on a thumb switch on the joystick controller used for flight control; "no" was indicated by moving the thumb switch backward, i.e., toward the subject. Reaction times were scored automatically to the nearest millisecond. If a subject did not respond within 2 seconds the trial was scored "no response."

Briggs and Swanson (1970) found a linear relationship between Sternberg task reaction time (RT) and central processing uncertainity (H_c) thus:

$$RT = a + b(H_c)$$

H_c values for this study are: 1.00, 1.50, 2.00 and 2.31 bits for the 1-, 2-, 4- and 6-item memory sets respectively. Because there is always a 2-choice response, response uncertainty (H_r) = 1.0 bit in each instance (Attneave, 1959).

In the equation for reaction time (RT), the intercept constant, a, reflects the time required for stimulus encoding, sampling and preprocessing at the input stage of human information processing plus the time to decode and execute a response in the output stage; the slope, b, reflects the time per test to complete stimulus classification functions at the central processing stage. Therefore, if a task performed *simultaneously with the Sternberg task, interferes with encoding or decoding processes, its effect* should be reflected by a change in the intercept value. Conversely, if the interference occurs at the central processing stage, the effect will be revealed by a change in the slope of the function.

Subjects. The data for this study were derived from the university students with an age range of 20-24 years. During the experiment a nominal cash incentive system was implemented to encourage performance. The amount of the incentive was based on the subject's relative standing in the group with respect to task performance criteria for each session. For dual task conditions the incentive value was weighted so as to emphasize priority for the flight control task,when it was present. The incentive was weighted in favor of the MFK task when it was paired with the Sternberg task.

Procedure. Prior to the experiment proper each subject was trained on all three tasks. Training sessions lasted two hours and were scheduled 2-4 times per week. Each subject was trained until task performance appeared to asymptote.

Each subject was tested under six different conditions: three single task conditions and three dual task conditions: Flight control, MFK and Sternberg choice-reaction task, alone; and flight control plus MFK, flight control plus Sternberg task and MFK plus Sternberg task. When the Sternberg task was combined with MFK it occurred only during periods when the subject was awaiting instruction for an MFK task of a given difficulty level. This was consistent with the interest in measuring cognitive loads associated with anticipation of MFK tasks rather than actual performance of them. The single task conditions preceded the dual task conditions for all subjects.

The type of maneuver "flown" was the independent variable for the flight control task. Although seven maneuvers were flown, preliminary analyses showed that not all maneuvers were discriminable in terms of the weighted tracking error scores. Hence, the maneuvers were combined into two groups labelled "easy" and "difficult". "Easy" maneuvers were climbing and diving turns. The error scores (X) were comprised as follows: X = (0.01) Δ altitude + (0.1) Δ airspeed for straight and level and stall, X = (0.01) Δ altitude + (0.1) Δ airspeed + (1.1) Δ g-load for straight and level turns, and X = (0.005) Δ vertical velocity + (0.1) Δ airspeed + (1.0) Δ g-load for turning dives and climbs. The delta values represent average error, i.e., deviation from the prescribed "fly to" value for the given flight parameter, per unit of time on the task. Altitude was measured in feet, airspeed in knots and vertical velocity in feet/minute. The flight parameter combinations and associated weights for each maneuver type were based on pilot opinion and research findings summarized in a separate report (Woodruff, 1972). MFK performance on multifunction keyboard tasks was measured in terms of task time and errors. The dependent measuare for the Sternberg task was reaction time. Errors and failures to respond within 2 seconds were also recorded.

RESULTS

Single Task Conditions. A simple analysis of variance (repeated-measures design) was applied to the scores for the flight control single task condition.The difference between easy and difficult conditions was statistically significant ($p < .05$). The mean and standard deviation for the easy condition were 1.09 and 0.17. Corresponding values for the difficult conditions were 5.11 and 1.51.

In order to obtain normality of distributions, MFK task times were subjected to reciprocal tramsformation prior to the analysis of variance. The effect of task difficulty proved significant statistically ($p < .001$) The Tukey hsd test showed that the mean for each task difficulty level differed from every other mean. Mean task times (seconds) and standard deviations (in parentheses) for the four difficulty levels were: I-3.97 (0.32); II-5.95 (0.53), III-7.43 (0.68), IV-9.87 (0.83). The average rate of information transmission via the MFK system varied from 1.8 bits/sec. to 2.6 bits/sec. across the four levels of MFK task difficulty. The switch action rate was slightly greater than 1 per second on the average. The method of least squares was used to fit a straight line to the baseline Sternberg data. The result is reflected by the following regression equation:

$$RT = 549 + 118(H_c)$$

Dual Task Conditions. Although mean flight control error was greater when the flight control task was combined with MFK tasks, the differences were not statistically significant. Similarly, MFK task times increased under dual task conditions, but the differences were not statistically significant. Flight control error scores were virtually identical for flight control alone as compared to flight control with the Sternberg task. The Sternbergy task had no statistically significant impact on MFK task time.

The method of least squares was used to fit linear equations to Sternberg response time data for each dual task condition. This permits comparison of intercept and slope values with those obtained for the Sternberg task baseline condition, for the purpose of localizing divided attention effects within the four stage information processing model.

Preliminary analysis showed no significant differences between levels of MFK task difficulty in terms of slopes and intercepts. Hence, a single regression equation was derived for the combined MFK levels. Equations for the resultant three dual task conditions are as follows:

Sternberg with MFK "Rehearsal"	$RT = 617 + 118(H_c)$
Sternberg with Easy Flight Control	$RT = 694 + 98(H_c)$
Sternberg with Difficult Flight Control	$RT = 855 + 31(H_c)$

F-tests (Snedecor and Cockran, 1967) indicate that (1) slopes and intercepts for the flight control conditions differ significantly from those for the baseline condition, and (2) the intercept value varies significantly between the baseline and MFK implicit rehearsal condition.

DISCUSSION

Locus of Divided Attention Effects. Interpreted in the traditional manner, the preceding results indicate that the effect of MFK "implicit rehearsal" is in the input or output stage of information processing only. Following the empirical evidence and logic of Briggs, et al (1972), the effect is probably in the input stage. The difference in intercept values amounts to a 12% average increase in input-output time attributable to MFK "implicit rehearsal."

Active flight control, on the other hand, appears to impact both input and central processing as evidenced by differences from baseline in both intercept and slope values for the regression equation. Moreover, there is an increase in input-output time (28% and 55% for easy and difficult flight control, respectively) and an increase in central processing rate. The central processing rate for the baseline condition is 8.47 bits/sec. as compared to 10.20 bits/sec. and 32.26 bits/sec. for the easy and difficult flight control conditions respectively. This increase in central processing rate under the dual task condition is consistent with results obtained by Lyons and Briggs (Briggs, et al., 1972). It was attributed to the subject's conducting fewer or less complete tests of the probe stimulus under the greater loading conditions. This apparent switch in mode of operation in the central processing stage may prove to be a valuable aid to identification of significant workload changes.

Information Transmission. Variations in Sternberg task response accuracy suggested the appropriateness of further information analyses, i.e., calculation of the average amount of information transmitted (which would reflect *all* the data), including erroneous responses and no responses. These values for the baseline and two levels of each dual task condition are presented below.

AVERAGE INFORMATION TRANSMITTED (H_t) IN BITS FOR STERNBERG TASK

	H_c			
Condition	*1.00*	*1.50*	*2.00*	*2.31*
Baseline	.85	.85	.88	.41
Easy MFK	.86	.85	.94	.42
Difficult MFK	.84	.88	.86	.32
Easy Flight Control	.82	.77	.79	.27
Difficult Flight Control	.72	.72	.79	.26

These data clearly indicate that the 6-item memory set (H_c = 2.31 bits) produced an overload situation for every task condition.

Effective Uncertainity Reduction. Since perfect performance is represented by H_t = 1.00 bit in each instance, the above tabled values were taken to represent percentage of the information reduction task effectively accomplished by the subjects. An information reduction task is defined as one in which the amount of uncertainty associated with the response is less than that associated with the stimulus (Coombs, Dawes and Tversky, 1970). Thus, using the measures of central processing time, a set of "effective uncertainty reduction rates" were derived and are plotted graphically in Figure 1. Note the consistent increase in efficiency as H_c goes from 1.00 to 2.00 bits with the overload effect that H_c = 2.31 for all conditions. Further study of Figure 1 suggests that cognitive reserve capacity is reduce by 20, 31, 45 and 54 percent by the four primary tasks (easy MFK "rehearsal," difficult MFK "rehearsal," easy flight control and difficult flight control) respectively.

Implications for digital avionics. With regard to the design issue addressed by this study, it appears that the multifunction switch concept places demands on the operator which may detract from the value of digital processing capabilities in avionics systems. This is not to say that the concept is necessarily inefficient. It is simply that it involves the concentration of uncertainty, normally distributed among the various dedicated instrument control/display interfaces, at a single interface. Moreover, uncertainty which is normally removed via separate controls and displays for each subsystem/function has to be eliminated via keyboard actions on each occasion that the operator interacts with the multifunction system. While the digitally-based MFK system is relatively efficient in terms of action and information transmission rates, the tasks are generally more complex and take longer than corresponding ones for dedicated instruments. Alternate designs for the MFK interface should be evaluated.

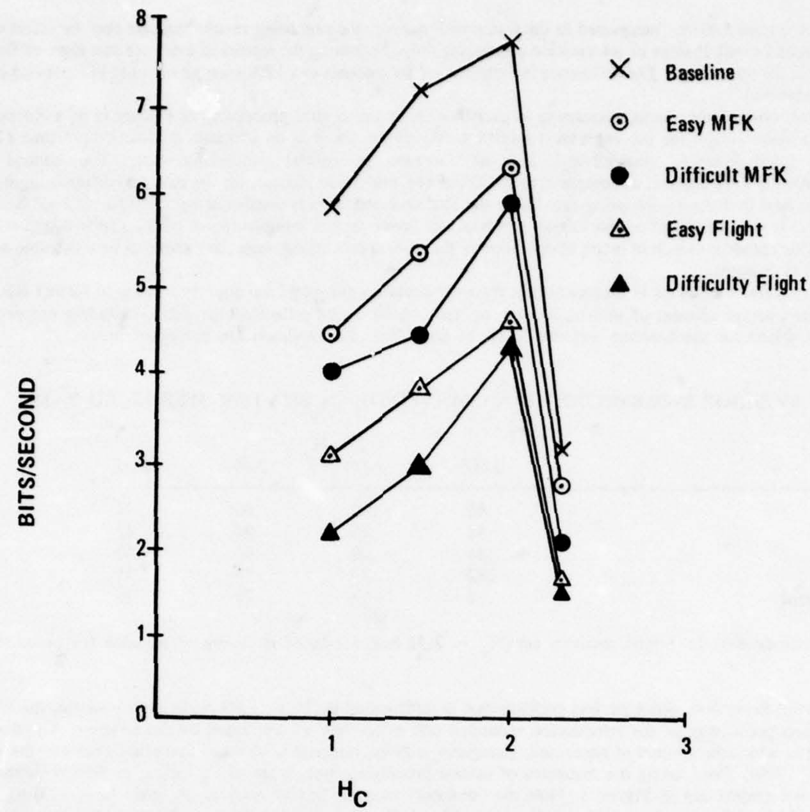


FIGURE 1. AVERAGE RATES OF EFFECTIVE INFORMATION REDUCTION ON STERNBERG TASK

REFERENCES

Attneave, F. *Applications of Information Theory to Psychology: A Summary of Basic Concepts, Methods and Results*. New York: Henry Holt and Company, 1959.

Brandt, W.E., Jr. and Wartluft, D.L. *Program documentation for the DAIS triple task experiment program*. Wright-Patterson AFB, Ohio: Aerospace Medical Research Laboratory, Technical Report AMRL-TR-75-24, December, 1975.

Briggs, G.E., Fisher, R.P., Greenberg, S.N., Lyons, J.J., Peters, G.L. and Shinar, D. *Multi-task time-sharing requirements*. Wright-Patterson AFB, Ohio: Aerospace Medical Research Laboratory, Technical Report AMRL-71-105, August, 1972.

Briggs, G.E., Peters, G.L. and Fisher, R.P. On the locus of the divided-attention effects. *Perception and Psychophysics*, 1972, *11* (4), 315-320.

Briggs, G.E. and Swanson, J.M. Encoding, decoding, and central functions in human information processing. *Journal of Experimental Psychology*, 1970, *86*, 296-308.

Coombs, G.H., Dawes, R.M., and Tversky, A. *Mathematical psychology*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1970.

Snedecor, G.W. and Cockran, W.G. *Statistical method*. Ames, Iowa. The Iowa State University Press, 1967.

Sternberg, S. High-speed scanning of human memory. *Science*, 1966, *153*, 652-654.

Sternberg, S. Memory scanning: mental processes revealed by reaction-time experiments. *American Scientist*, 1969, *57*, 421-457.

Swanson, J.M. and Briggs, G.E. Information processing as a function of speed versus accuracy. *Journal of Experimental Psychology*, 1969, *81*(2), 223-229

Woodruff, K.R. *Development of a rate-of-closure display using a CRT symbol generator*. Systems Research Laboratory, Dayton, Ohio: Systems Research Laboratories, Inc., Technical Report, September 1972.

Zipoy, D.R., Premselaar, S.J., Gargett, R.E., Belyea, I.L., and Hall, H.J., Jr. *Integrated information presentation and control system study*. Wright-Patterson AFB, Ohio: Air Force Flight Dynamics Laboratory, Technical Report AFFDL-TR-70-79, Vols I, II and III, August, 1970.

RESEARCH ON THE STRESSES OF WARTIME CAPTIVITY

Edna J. Hunter
Family Studies, Naval Health Research Center
and
Family Research Center, U.S. International University

A modicum of stress can be motivating; too much can be devastating. Interstingly, certain situations can create stress for a given individual, and yet a situation which is stressful for one individual may create no stress whatsoever for another. Moreover, a situation stressful on one occasion may not be so on another for the same individual. It has been noted that personal characteristics are related to one's appraisal of threat, which, in turn, determines how stressful any situation is perceived to be (Lazarus, 1966). Thus personality factors, as well as prior experiences, enter into the equation for determining stress level and also one's response to a particular stressor.

For almost twenty years researchers at the Naval Health Research Center (NHRC) in San Diego, California, have been studying how men (and more recently, women) cope with stressful environments, e.g. the stresses encountered working within the confines of ships at sea, wintering over in the Antarctica, living beneath the sea, etc. It was logical, then, when the Center for Prisoner of War Studies (CPWS) was established by the Navy Bureau of Medicine and Surgery, to place it under the auspices of the Naval Health Research Center.[1].

The Center for Prisoner of War Studies, unlike other divisions within NHRC, is a collaborative Army-Navy-Marine Corps activity, jointly funded by the Navy Bureau of Medicine and Surgery and the Office of the Army Surgeon General in 1972 to monitor the physical, psychological, social and family adjustments of the 241 Army, Navy and Marine Corps prisoners of war (POWs) returned from Southeast Asia in 1973, as well as the adjustments of their families and the families of servicemen missing in action (MIA). This prospective longitudinal study comes under the umbrella of preventive medicine, with the research mission merely an adjunct to the primary mission of making certain the POWs were brought back in the best possible fashion, medically speaking, given good basic medical evaluations upon return and reexamined annually for several years post-release. These actions were designed to ameliorate or possibly even avoid entirely, the anticipated long term residuals of captivity stress.

Stated succinctly, the *research* mission was "the achievement of a better understanding of the multidimensional impact of captivity upon men and their families." A review of the literature shows that follow-up studies of concentration camp victims and American POWs of the Japanese, North Koreans and North Vietnamese indicate that permanent psychic and psychophysiological damage can indeed occur to adult human beings if they are subjected to prolonged malignant and cataclysmic stress (Hunter, 1978). In other words, the extraordinary stresses of incarceration are related to a heightened vulnerability to physical and psychological health problems over time (Segal, Hunter & Segal, 1976), and the onset of symptoms may even occur for the first time several years *subsequent* to release from captivity.

Payoffs from this longitudinal study of a discretely defined population which had experienced stresses impossible to replicate in the laboratory, have implications which extend far beyong the immediate accumulation of knowledge about the limited populations being studied; e.g., the reexamination of the The Military Code of Conduct and its interpretation, modifications in survival training, improved selection of personnel for high-risk assignments, and the development of measurement techniques for assessing family functioning, etc. Perhaps it should be noted that these in-depth studies of military families are a research "first" which can perhaps provide insights into family dynamics which will result in a better understanding of military families generally––happier, healthier families with higher morale, more efficient performance on the part of the service member, and decreased cost requirements for health-care delivery services.

Many questions have been posed by our study such as:

- What factors determine who dies and who survives captivity?
- Are there differences in the later health and adjustment of former prisoners which are related to length of *imprisonment*, conditions of captivity, time spent in solitary confinement, men's perceptions of the stresses of torture, etc.?
- Are there significant differences in the later health and adjustment of officer versus enlisted men which might reflect differences in coping abilities during captivity?
- Does the stress of captivity have a cumulative effect, or perhaps an accelerating one, so that in future years the returned POWs will pay an additional price for the years of incarceration they endured?
- Do the members of the families also show heightened vulnerability to psychological and physical illness which related to the social and emotional stresses they endured while their husbands or fathers were incarcerated? (Hunter & Plag, 1977).

These are but some of the specific hypotheses we have been examining during the past five years. We have already answered some questions; others remain to be answered through future study. The staff of the Family Studies Branch of CPWS began interviewing the families in 1971, prior to the men's return. We are at the moment in the process of personally interviewing the families for the sixth and probably final time; the family follow-up will end in September 1978. The physical examinations for the men, however, will continue through 1982––at least for the Navy and Marine returned POWs. Medical examinations are also being done annually at the Navy Aerospace Medical Research Laboratory (NAMRL) on a carefully matched control group. With such a comparison group, we shall be better able to determine which findings are attributable to captivity stresses.

I shall not go into detail on our findings, but because of its relevance to our topic today, I want to mention that one of the main findings emanating from these studies is that *each POW perceived his situation in a unique manner.* This difference in the perception of the captivity experience was a function of his personality, his experience with stress, and his characteristic way of interacting with his environment. Personality characteristics were also related to the firmness of resistance the man demonstrated during captivity. Certain personality traits, in fact, resulted in the POW receiving harsher treatment by the captor.

---

[1]The Naval Health Research Center was previously known as the Navy Medical Neuropsychiatric Research Unit until the activity's functions were expanded in 1974 to include immunological and toxicological studies.

Related ongoing research at the Center has shown that a man's locus of control—that is, whether he perceives his reinforcement as being the product of his own efforts (Internals) or under the influence of fate or others (Externals)—has a significant relationship with performance in a coercive environment. We must remember that we are dealing with the *survivors* of the captivity experience. In viewing the returnees as survivors, we can expect to learn something about how they perceived the stress, how they *coped* with stress, what price they have *paid*, and *may* pay in the furture in terms of physical and psychological health. And finally, we hope to learn what they may have *gained* from the experience in terms of greater appreciation for life and family, increased insight into their own values systems and their ability to cope with extreme stress (Richlin, 1977).

The families, too, experienced their own form of "captivity," and we have been looking at how wives and children coped with husband-father absence, and what the effects of family separations are on the personal adjustment and health of family members. Our in-depth longitudinal study of this small group of men and families who have experienced prolonged stress offers a unique opportunity to understand how families cope in unique ways.

Several of our studies examined the strategies used by men and their families in coping with stress (Deaton, Berg, Richlin & Litrownik, 1977; McCubbin, H.; Dahl, B.; Lester, G.; Benson, D. & Robertson, M., 1976). We found that coping activities during solitary confinement developed according to a definite time pattern. Initially concerned with family and the past, later the POW developed coping activities related to self-development. The use of self-development strategies appeared to be inversely related to the degree of the POW's physical deprivation and the perceived stress level (Richlin, 1977). Wives also used a variety of coping styles—some functional and others dysfunctional—in dealing with family disruption (McCubbin et al., 1976). Our studies show that maintenance of a husband/father role within the family during the captivity period was an important factor for successful family reintegration following return (McCubbin, Dahl, Lester & Ross, 1975). Ironically, successful coping for the wife during the separation period had required at least a partial "closing out" of the father's role.

Other preliminary studies carried out at CPWS have shown that (1) the amount of solitary confinement the POW experienced during captivity was highly related to his perception of how well he was doing in his career three years subsequent to return; (2) POW families, as a group, were more "female-centered" or matriarchal than the more "traditional" control families; (3) poorer father-child relations were related to longer captivity duration, more abuse in captivity, and firmer resistance toward the captor.

Indeed, both the events of casualty and reunion have been shown to have been stressful family crises which called for extreme amounts of adjustment on the part of family members, which in most cases could be considered to be merely *normal* adjustment to a very *abnormal* situation.

In closing, I quote one physician who has monitored the POWs closely throughout the post-captivity period: "It is now apparent that the process of recovery from the stress of shootdown, capture, captivity, and repatriation appear to require, among other things, recovery of self-esteem through reintegration with the group; the POW group, the military, the family; and society . . . . To the degree that there is failure, there will be . . .psychopathology." Although the studies at the Center may at times appear to focus heavily on psychopathology, they have also afforded new insights both into the manner whereby POWs have coped with tremendous stresses and into the ways in which the ordeal of captivity served to strengthen them and build new resources—both for the men and their families (Segal, Hunter & Segal, 1976).

REFERENCES

Deaton, J.; Berg, S.; Richlin, M. & Letrownik, A. Coping activities in solitary confinement of U.S. Navy POWs in Vietnem, *J. of Applied & Soc. Pshychol.*, 7(3), 1977, 239-257.

Hunter, E.J. The Vietnam POW Veteran: Immediate and long term effects of captivity. In Figley, C.R. (Ed.), *Stress disorders among Vietnam veterans: theory, research and treatment implications.* New York: Brunner/Mazel, 1978.

Hunter, E.J. & Plag, J.A. The longitudinal studies of prisoners of war and their families: An overview. *NHRC Technical Report No. 77-26*, 1977.

Lazarus, R.S. *Psychological stress and the coping process*, New York: McGraw-Hill, 1966, 210-257.

McCubbin, H.; Dahl, B.; Lester, G.; Benson, D. & Robertson, M. Coping repertoires of families adapting to prolonged war-induced separations. *J. of Marr. & Family*, Aug., 1976, 461-471.

McCubbin, H.; Dahl, B.; Lester, G. & Ross, B. The returned prisoner of war: Factors in family reintegration. *J. of Marr. & Family*, Aug., 1975, 471-478.

Richlin, M. Positive and negative residuals of prolonged stress. *NHRC Technical Report No. 77-29*, 1977.

Segal, J.; Hunter, E.J. & Segal, Z. Universal consequences of captivity: Stress reactions among divergent populations of prisoners of war and their families. *Int. J. Soc.,Sc.*, Vol. XXVIII (3), 1976, 598-609.

DOD RESEARCH ON PSYCHOLOGICAL STRESS

James Lester
U.S. Office of Naval Research, Boston Branch

I. Introduction

This paper is based on an intensive look at all the abstracts obtained from the Defense Documentation Center by requesting the Center to print out all the information it has on reports relating to psychological stress. This request brought me 566 abstracts, some of which I had to exclude from my survey because they were mistakes, because they did not fit the notion of psychological stress, or because they were not acutally done nor funded by DOD. The remaining abstracts (and not the original reports or books) were considered in some detail and from several points of view as you will see.

My purpose in reporting on this is very broad and general; it is *not* to achieve an in-depth survey and critique of any particular kind of stress research. The question I have tried to answer has been something like, "What work has the DOD sponsored on the topic of psychological stress?" I'm going to talk about the *extent* and the *focus* of DOD-sponsored stress research, and then discuss one or two specific content domains in a little more detail. To put this another way, I'll not be talking about behavior under stress, but rather about the behavior of stress researchers; where have they invested their effort? As an index of the investment of effort I have used a simple tally of number of reports; perhaps number of pages would have been better, but I had no help and did everything by hand, and simply didn't feel like putting out the extra effort that counting pages instead of reports would have required. I don't know how well the number of reports written really measures the extent of DOD interest—the phrase "DOD interest" is a convenient shorthand for what I'm talking about and one I hope will catch your interest and focus your attention. But what I'm really talking about as raw data here is the number of reports of various kinds put out under DOD sponsorship.

II. Extent of DOD Interest in Stress

The earliest report in this data base is from 1951, and the most recent is early 1977. In that period 344 reports appeared describing studies or wisdom paid for by the DOD. In-house DOD laboratories and contractors contributed about equally to this total—47% and 53% respectively. Incidentally, the DDC printout indicates for each report whether or not it is also available in a professional journal; among the 25% published in the psychological literature there was an *im*balance between in-house and contractor reports, with more of the published subset coming from in-house labs (65%) than from contractors.

There were 161 in-house reports in the period 51-77 and I'd like to move in for a close-up of that lot now. And I'd like to ignore for a moment the 12 reports appearing after the end of 1974, so as to deal with five equal time periods of five years each. Slide 1 shows the number of reports appearing each year, and it is clear from this that DOD interest in general increased. Forty percent of the total were produced in the most recent of the five periods, as shown in Slide 2 which summarizes report production by five-year periods.

Of the 149 in-house reports we're now looking at, Navy labs provided 55%, Army labs provided 25%, and Air Force labs 19%. But these relative proportions were different in the various time periods. The fact is that from 51 through 74 Navy labs steadily increased their relative contribution, while Army and Air Force labs showed a general decline in *their* (relative) contribution. With regard to stress, both the Army and Air Force hit their peaks of report productivity in the early '60s, the Navy did so in the early '70s. Getting ahead of the story a bit, there is a subject matter difference associated with these productivity peaks; in the '60s (but not so much later) the Army was heavily concerned with the effect of various stressors on weapon systems performance (especially the effects of noise, restricted environments, temperature, and vibration), while the Air Force was specializing in studies of confinement and sensory deprivation; in the '70s the Navy has come forward with a large number of studies of life stress as it relates to health. In the period 70-74, Navy labs contributed 73% of the in-house DOD stress-related reports, while the other services contributed 14% each.

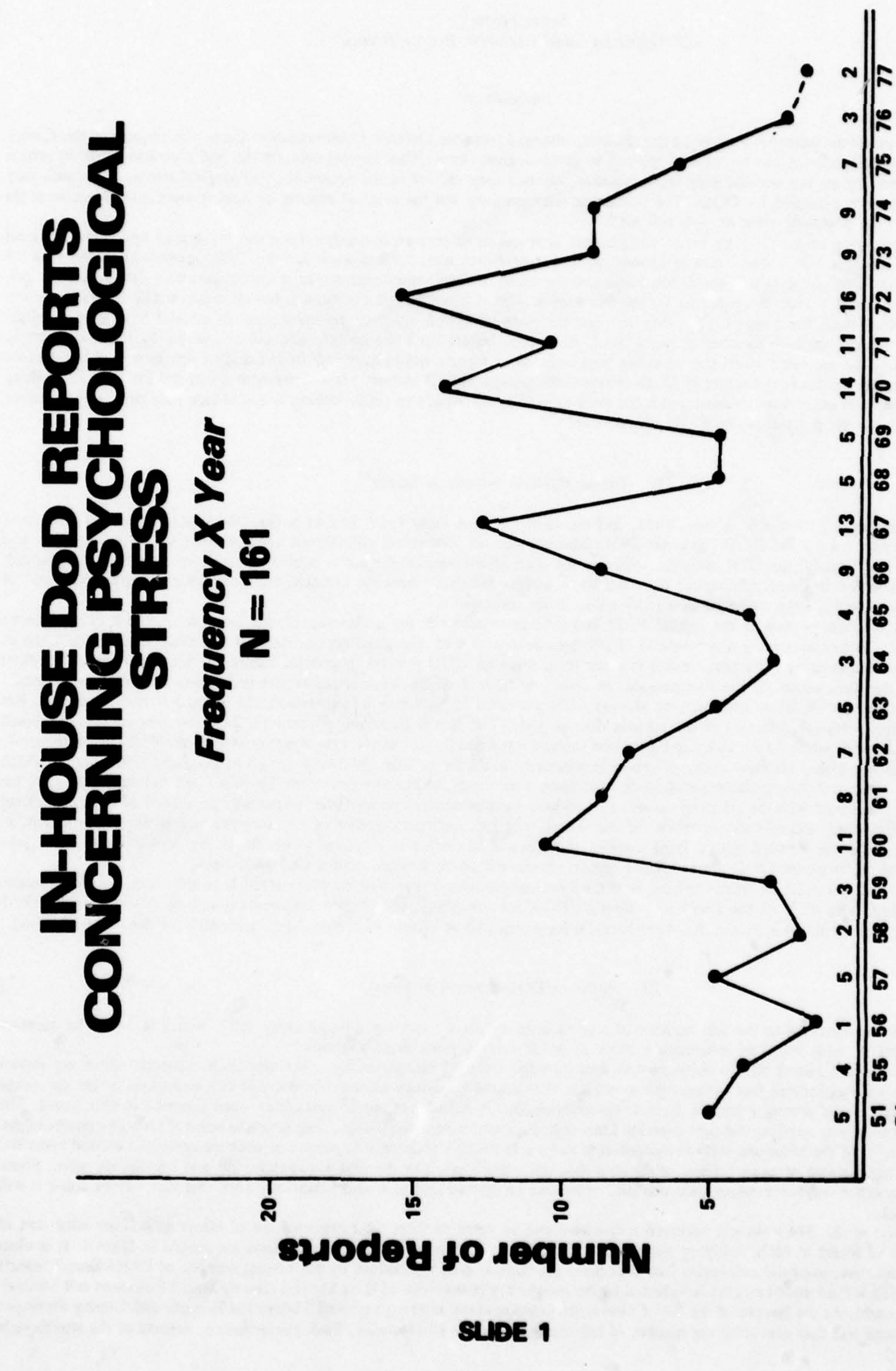
For each service one particular lab seems to have been the stress-specializing center. For the Navy it has been the Naval Health Research Center in San Diego, with 49% of the Navy total through 1974; for the Army, the Human Engineering Lab at Aberdeen, with 30% of the Army total; and for the Air Froce, the Aero-Medical Research Lab at Wright-Patterson AFB, with 48% of the Air Force total.

III. Focus of DOD Interest in Stress

In this section I am returning to the full number of reports cited by DDC, running through early 1977, which is 344. The question I'm addressing here is, what kinds of information about stress do these reports make available?

To answer this question I read all the abstracts and then invented some 15 categories for describing their contents, which are shown in Slide 3. These were categories that (1) seemed to capture the natural groupings among reports, and (2) seemed to reflect the range of interests that might exist among potential users of the information. In actual fact the 15 categories were invented in two stages. The first nine emerged mainly by applying the first criterion I mentioned: natural content groupings, seen from the point of view of a psychologist. But as I progressed with the laborious task of categorizing these references I realized that particular user interests (the second criterion) might require the highlighting of some aspects of the data base not made explicit by the first nine categories, and I added six more. There are no doubt many other ways one might slice this pie, depending on special interests, and I can only hope that this way of doing it will be generally useful.

In any case, each of the 344 abstracts received a check in one or more of these 15 categories, in an effort to indicate what sort of information might be found in each. Tallying the number of checks in each category gave the information shown in Slide 4. It is clear that among the basic psychological categories (the first nine) the content area touched on by the largest number of DOD-funded reports (38% , or 130) is Task Performance. This is followed by the category reviews (with 26% or 88) and then by Social Processes and Methodology (with 17% each). At the bottom of the list of nine is information about Interventions and Theory (04% each) and Coping Strategies (03%)—this last with less than one-tenth the number of reports accorded the front-runner, Task Performance. Several of the afterthought

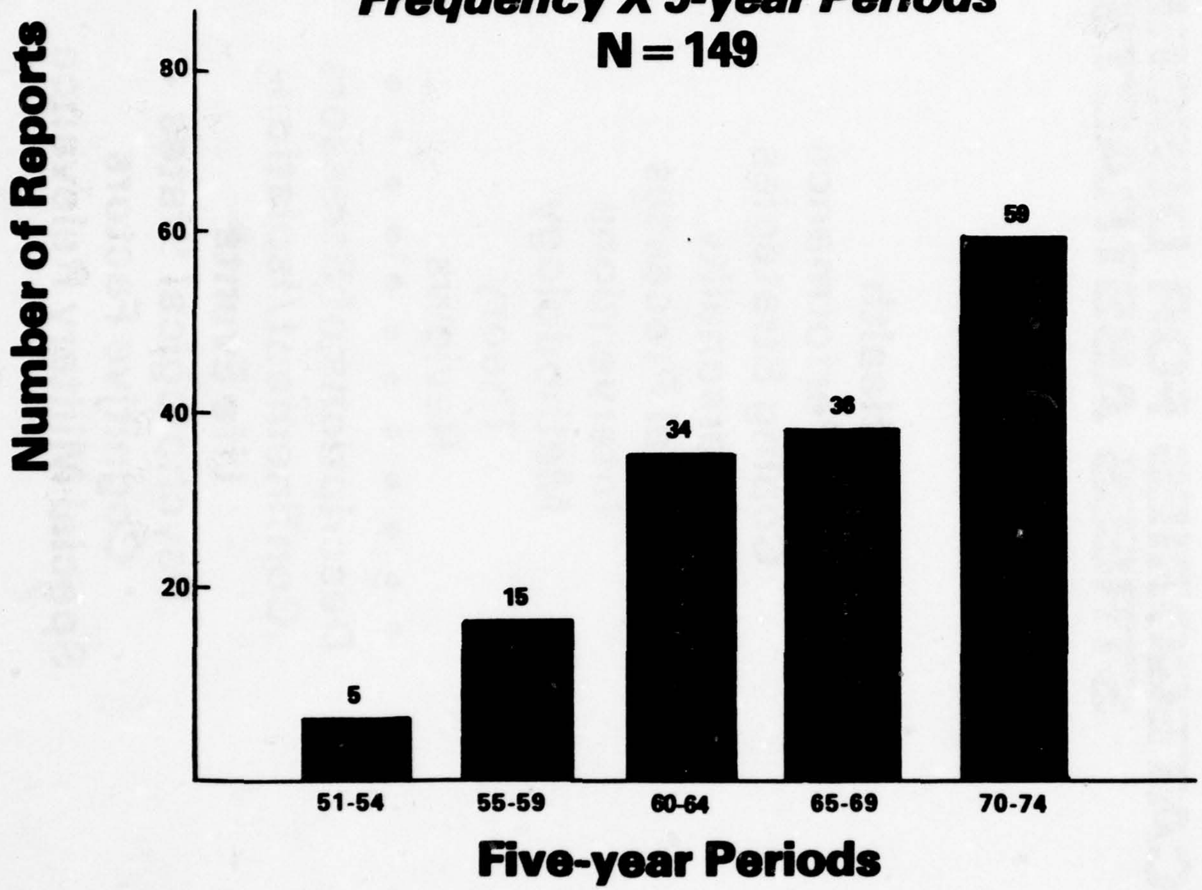

IN-HOUSE DoD REPORTS CONCERNING PSYCHOLOGICAL STRESS
Frequency X Year
N = 161
Number of Reports
20
15
10
5
5 4 1 5 2 3 11 8 7 5 3 4 9 13 5 5 14 11 16 9 9 7 3 2
51-54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77
Years

SLIDE 1

IN-HOUSE DoD REPORTS CONCERNING PSYCHOLOGICAL STRESS

Frequency X 5-year Periods

N = 149



SLIDE 2

CATEGORIES FOR DESCRIBING STRESS ABSTRACTS

Health
Task Performance
Coping Strategies
Personality
Social Processes
Interventions
Methodology
Theory
Reviews

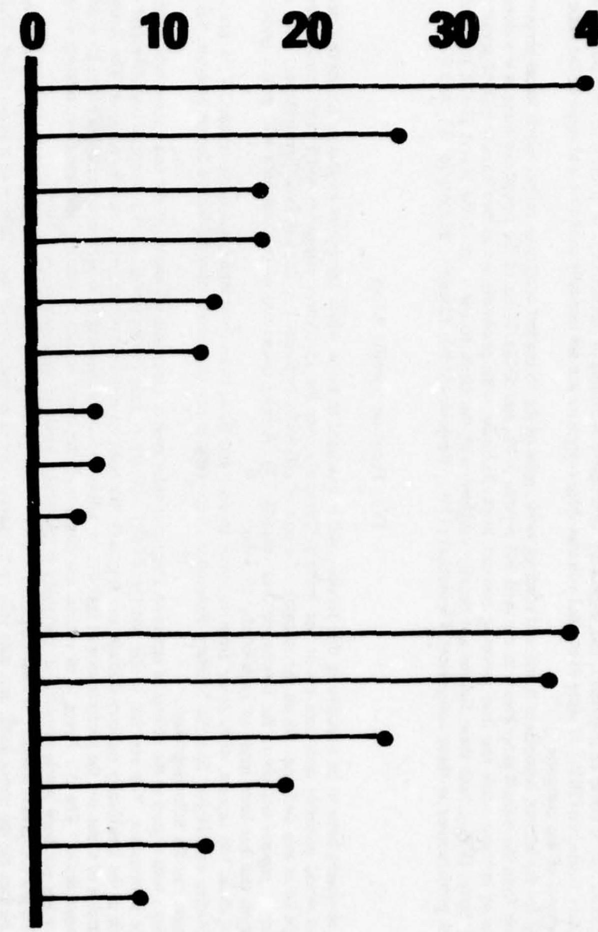
• • • • • • • • • • • •

Descriptions of Stressors
Confinement/Isolation
Life Events
Psychological States
Cognitive Factors
Special Military Relevance

SLIDE 3

PERCENTAGE OF REPORTS IN EACH CATEGORY

Category	%
Task Performance	38
Reviews	26
Methodology	17
Social Processes	17
Personality	14
Health	13
Interventions	4
Theory	4
Coping Strategies	3
Descriptions of Stressors	37
Special Military Relevance	36
Confinement/Isolation	25
Cognitive Factors	18
Psychological States	11
Life Events	7



SLIDE 4

categories turned out to be important, too, especially Descriptive Accounts of Stressors (37% of the total) and Studies with Highly Specific Military Reference (36%). (I want to remind you that these categories are not mutually exclusive and some of the same reports will appear in several of the categories.)

As for the balance between in-house and contract work, most of the 15 content categories reflect nearly equal proportions, as you would expect from the overall proportions in the total 344 reports (47% and 53%). In six of the categories there is a noticeable *im*balance, two favoring in-house work and four favoring contract work. In-house work predominates in two closely related areas: Health (72%), and Life Stress (87%); both these figures are heavily weighted with the fairly recent work of the Naval Health Research Center. Contract work predominates in these categories: Theoretical (77%), Reviews (65%), Cognitive Factors (66%), and Coping Strategies (64%).

IV. Focus on Specific Areas

I obviously have to be selective in this section, and it was hard to know which categories to choose for more detailed attention. Each of us would probably distribute his or her interest differently over the 15 content categories and I can't guess where the most interest might lie in this audience. So on the strength of a sort of efficiency criterion I've chosen Task Performance, since a discussion of that group or reports will cover the largest number, namely 130. A brief discussion of this one area will give a good idea of the kind of analysis that has been made of each of the 15 areas.

Of these 130 reports, 47% come from in-house sources and 53% from contractors. Of these reports, 42% are also cross-classified as Studies with Highly Specific Military Reference. According to my subjective, nonquantitative factor analysis, the 130 reports can be clustered in the following ways.

First, with regard to the effects of stress on some particular type of task; slides 5 and 6 list the tasks about which this data base contains some information. You see that in the not specifically military domain the emphasis is on perceptual cognitive and psychomotor tasks, while in the specifically military domain the emphasis has been on air crew performance, mainly piloting. The Army has done a series of studies of stress and the performance of tank crews, a few on factors affecting rifle performance, and a series on what they call "combat-relevant skills." The Air Force is of course associated with the many studies of piloting and questions relating to space flight, and the Navy has not been much involved in the study of particular kinds of tasks (so far as I know).

The just-mentioned studies put their emphasis mainly on task performance and how it is affected by stress; another main body of studies emphasizes the the stress itself, and uses task performance only as an indicator of the existence of a stress effect (Slide 7). Among DOD studies of this second kind, 50% have to do with stressors in the physical environment, and here the great bulk have focused on noise as a stressor; heat, vibration, and flickering lights have also received some attention. The next largest category of stressors is confinement or isolation, and these I believe have been mainly oriented toward understanding the problems of the POW, the astronaut, or the tank crew. Cognitive stress—such as the effect of a speed requirement on a mental task, or overload in a cognitive or perceptual performance—runs a weak third, after environmental stressors and social isolation. The last cluster of any significance (and it is a small one) concerns factors that tend to counteract the effects of stress on performance, such as the provision of feedback, attitude and motivation, personal background factors, and drug effects.

Still in the Task Performance area, 10 reports are cross-classified as mainly concerned with Methodology (ranging from a 1953 "Bibliography for the development of stress-sensitive tests" to a 1973 report on ". . .application of identification methods for conditions"). Another 10 reports can be characterized as being more general than the ones described up to now—that is, they deal with "human behavior" under stress, or consider a mixed variety of behaviors together. These 10 range from a 1951 "Review of research on the effects of psychological stress on performance" to a 1975 report on the "Measurement and prediction of operational fatigue."

Let me finish with remarks about one other category among the 15 main content areas: Reviews. This will give a general idea about how recently and with what focus DOD reports have surveyed large segments of the field. One-quarter (or 88) of the 344 reports were tallied as giving information of a review type, and once again the in-house vs contract proportions were 47 and 53% respectively. With regard to domains covered in these reviews, the largest number concern isolation or confinement (nearly 30%); of these the majority review studies of experimental isolation (as opposed camp). It's not surprising that a fairly large proportion (14%) deal with performance of some kind, since Task Performance is so heavily represented in the total. It *is* somewhat surprising that Social Processes is equally well represented, with Health and Well-Being also at about the same level.

Looking at Reviews in a different way (Slide 9), one sees that 11% of the reports are essentially bibliographies; four of these appeared in the '50s, the other six in the '60s, with the most recent bibliography dated 1965. A somehwhat more demanding genre, the Literature Review or Survey also contains 11%, with three in the '50s, six in the '60s, and only one in the '70s (on operational fatigue, 1975). Finally, six of the eighty-eight reports are essentially proceedings of conferences, three in the '50s and three in the '60s. I know there has been no shortage of conferences on stress more recently, but the initiative for them has apparently moved out of DOD.

V. Conclusion

This set of reports is rich in variety and probably varies all over the map in quality. One might say it is a quilt made up of many patches, and I'm not sure how much I would depend on it for warmth. I do have the impression that what I said about conferences applies to the subject as a whole: The DOD probably played a major role in providing research on stress in the '50s and '60s, but I suspect the rate of production is declining, while civilian investigators and non-DOD sources are producing an increasing number of studies. I have in hand a set of 800 abstracts from the psychological literature, produced for me by the Psychological Abstracts Search and Retrieval service, on which I will be carrying out analyses similar to these, so that I can also look at what has been going on outside of DOD with regard to stress.

There is obviously no way to summarize either what I've said, or the larger domain of things I might have said but didn't. I am continuing to zoom in on specific areas and to try to describe what kind of information might be found there, and I intend to make all of this available in the form of at least a 345th DOD report, possibly more than one.

STUDIES EMPHASIZING EFFECTS RATHER THAN CAUSES

Cluster	Number of Reports
Not Specifically Military	31
• cognitive/intellectual tasks	15
• psychomotor tasks	7
• verbal performance	4
• group creativity	3
• man-machine performance	1
• shelter management behavior	1

SLIDE 5

STUDIES EMPHASIZING EFFECTS RATHER THAN CAUSES

Cluster	*Number of Reports*
Specifically Military	36
• aircrew performance	17
• space flight	5
• combat-relevant skills	5
• extended operations	4
• rifle performance	3
• submarine crew effectiveness	1
• mental health technicians	1

SLIDE 6

STUDIES EMPHASIZING CAUSES RATHER THAN EFFECTS

Cluster	*Number of Reports*
• Environmental Stress	25
• Confinement/Isolation	13
• Cognitive Stress	6
• Miscellaneous	7
Stress Counteractants	6

SLIDE 7

REVIEW CATEGORY: DOMAINS COVERED

Domain	*Number of Reports*
• Isolation, confinement	26
• Performance	12
• Social Processes	11
• Health, Well-being	10
• Space	4
• Task-induced Stress	2
• Extreme Environments	2
• Combined or General Stresses	20

SLIDE 8

REVIEW CATEGORY: TYPE OF REVIEW

			Number of Reports
• Bibliographies			10
	1950's	(4)	
	1960's	(6)	
	1970's	—	
• Literature Reviews			10
	1950's	(3)	
	1960's	(6)	
	1970's	(1)	
• Symposia, Conferences, *etc.*			6
	1950's	(3)	
	1960's	(3)	
	1970's	—	

1950's — 10
1960's — 15
1970's — 1

SLIDE 9

A COMPARISON OF THREE TREATMENTS FOR TEST ANXIETY

Richard L. Hughes
U.S. Air Force Academy

Text anxiety is a situation-specific form of stress that affects as much as fifteen percent of college students. It leads to imparied performance on certain sorts of tasks and it is subjectively unpleasant.

The sorts of tasks most adversely affected by test anxiety are those involving demand for high achievement and the evaluation of performance (Spielberger, Anton & Bedell, 1976). Since such conditions characterize the atmosphere surrounding most college academic evaluations, text-anxious college students are at a severe disadvantage. Such students may fare poorly academically, not from lack of effort or intellectual competence, but because of debilitating anxiety on tests. It is therefore encouraging that numerous studies have demonstrated significant reductions in test anxiety following relatively brief behavioral treatments. This study reports a comparison of several different treatments for reducing test anxiety in a college population.

Various techniques have been used in treating test anxiety, but most are based on the assumption that reducing emotional arousal is the essential factor in treatment. Systematic desensitization, with its emphasis on the cultivation of a deeply relaxed musculature, has been the single most popular treatment. It is interesting to note, though, that despite the conceptual centrality of muscle relaxation to most forms of treatment, few studies have assessed actual changes in muscle tension following treatment. This would seem to be important, since some work with electromyographic biofeedback (Reinking & Kohl, 1975) suggests that standard procedures of training deep muscle relaxation may be relatively ineffective.

Another criticism of desensitization is that it focuses too much on the emotional component of test anxiety and not enough on the cognitive or worry component (Liebbrt & Morris, 1967). Wine (1971), among others, has argued persuasively that test anxiety is a cognitive problem more than an emotional one, and hence, that a more cognivitely oriented treatment is appropriate. This could include, for example, a program which fosters more constructive attitudes regarding academic evaluations.

The present study was designed to test two particular questions which emerge from the preceding discussion. First, is electromyographic biofeedback superior to verbal procedures for training deep relaxation in treating test anxiety? Secondly, is it effective in such treatment to present positive attitudes toward test-taking during periods of deep muscle relaxation?

METHOD

Subjects

The Test Anxiety Scale (I. Sarason, 1972) was administered to all sophomore cadets (approximate 300) in an undergraduate psychology course. Cadets who reported high test anxiety were invited to participate in a program to reduce test anxiety. There were 26 cadets who volunteered and they were randomly assigned to 1 of 3 treatment conditions. The experimental groups were: biofeedback relaxation, biofeedback desensitization, and traditional desensitization. The entire program lasted approximately three months. *Ss* reported individually for baseline and treatment sessions on a fairly regular basis averaging one session per week. All baseline and treatment sessions were completed during one academic semester except in the case of one *S* who did not complete the physiological post-testing until after his final examinations.

Treatment Conditions

Biofeedback relaxation. After receiving a brief introduction to biofeedback theory and procedure, *Ss* received four twenty-minute sessions of electromyographic feedback of frontalis muscle tension. *Ss* received both auditory and visual feedback. Then for two twenty-minute sessions *Ss* practiced tape-recorded relaxation instructions. During the seventh training session *Ss* relaxed for several minutes, after which they listened to a tape recording of positively-worded statements about testing (e.g., "I will concentrate effectively").

Biofeedback desensitization. After receiving a brief introduction to biofeedback theory and procedure, *Ss* received four twenty-minute sessions of electromyographic feedback of frontalis muscle tension. During the next three twenty-minute sessions *Ss* listened to a standardized desensitization tape depicting several extremely stressful testing situations.

Traditional desensitization. This group received identical treatment as the biofeedback desensitization group, except that relaxation was trained via standard verbal procedures rather than with biofeedback. During these relaxation training sessions *Ss* listened to different tape recordings which presented both progressive relaxation and autogenic training exercises.

Dependent Variables

All *Ss* participated in eleven sessions. The first two and last two sessions were for pre and post-testing. During these sessions *Ss* completed the comprehension portion of the Nelson-Denny Reading Test, Forms C and D, and equivalent forms of a locally-generated test composed of analogies and algebraic word-problems. The order of presentation of these tests was counterbalanced. Frontalis muscle tension was assessed using a Bio Feedback Systems, Inc. B-1 electromyographic biofeedback instrument (though *Ss* did not receive feedback during the assessment sessions). Frontalis tension was recorded while *Ss* listened to tape-recorded descriptions of academic testing situations and imagined themsleves in the situations. Physiological readings were taken while the *S* was seated in a soundproof room with a soft level of background white noise. The Test Anxiety Scale was readministered after the final physiological assessment period.

At the beginning and end of each of the seven training sessions each *S* also rated his or her subjective level of tension. *Ss* were asked, "On a scale of one to ten, ten being as tense as you have ever felt, how do you feel now?"

The final dependent variable was semester grade point average for the semester preceding and the semester of the study.

Counselors

Seven counselors with varying degrees of counseling experience participated in the study. Three were professional counselors with extensive counseling experience and four were undergraduates with some counseling experience. Each counselor worked with the same Ss across treatment sessions and each counselor was responsible for *Ss* from all treatment conditions.

RESULTS

Results of this study are summarized here:

1. All treatments significantly reduced test anxiety. There was no difference between treatments.

2. Treatments showed little effect on frontalis muscle tension, but this was due primarily to very low levels of frontalis tension during pretesting. In other words, *Ss* were at relatively relaxed levels of muscular tension before treatment, despite self-reports of considerable anxiety. Thus, there was little range for possible modification of scores. This physiological data probably does not reflect an ineffectualness of the treatments so much as it may an inappropriateness of them.

3. Treatments had minimal effects on several performance measures. There were no significant differences in pre and post-measures of grade point average, reading, or other cognitive tasks.

DISCUSSION

Despite the popularity of relaxation procedures and specifically desensitization for reducing test anxiety, this study raises several questions about the adequacy of such treatments. While all treatments in this study did, in fact, reduce test anxiety, this did not appear to be due to actual changes in the ability to relax since pre and post-levels of muscular tension were not substantially different. Thus, the theoretical basis for relaxation-type treatments is called into question by the absence of results demonstrating high levels of physiological tension before treatment. This is not a completely unexpected finding since Wine previously suggested that emotional arousal is not, after all, the critical feature of test anxiety. In that sense, perhaps anxiety is an unfortunate word to describe this specific problem, since the word anxiety clearly implies physiological symptoms. Furthermore, perhaps it is an inappropriate focus on the emotional rather than the cognitive aspects of test anxiety that lay behind the absence of any treatment effet on performance measures. It should be noted that this is a common finding in similar studies. Finally, while it may be true that cadets constitute an untypical group of college students, it is nevertheless clear that their test anxiety scores were comparable to those of students who receive treatment for test anxiety at other institutions. The cadets *did* constitute a test-anxious group.

REFERENCES

Liebert, R.M. & Morris, L.W. Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, 1967, *20*, 975-978.

Reinding, R.H. & Kohl, M. Effects of various forms of relaxation training on physiological and self-report measures of relaxation. *Journal of Consulting and Clinical Psychology*, 1975, *43*, 595-600.

Sarason, I.G. Experimental approaches to test anxiety: Attention and the uses of information. In C.D. Spielberger (Ed.) *Anxiety: Correct Trends in Theory and Reason*, Vol 2, New York: Academic Press, 1972.

Spielberger, C.D., Anton, U.D., & Bedell, J. The nature and treatment of test anxiety. In M. Zuckelman & C.D. Spielberger (Eds.) Emotions and Anxiety: New Concepts, Methods,and and Applications. New York: LEA Wiley, 1976.

Wine, J. Test anxiety and direction of attention. *Psychological Bulletin*, 1971, *76*, 92-104.

STRESS, PERSONALITY, AND ILLNESS IN A GROUP OF ARMY OFFICERS

Suzanne C. Kobasa
University of Chicago

In the past twenty years, an exceptional number of research studies conducted by physicians and social scientists at several research centers and universities have been dedicated to proving *that stressful life events cause illness* (cf. Dohrenwend & Dohrenwend, 1974; Gunderson & Rahe, 1974). Both single, severe stresses and multiple, moderate stresses have been implicated in the onset of symptomatology. Any life event which can be characterized as causing change or demanding readjustment in an *average person's* normal routine qualifies as a stress, and is thereby suspected of having etiological significance.

The typical stress and illness study is correlational and retrospective. A large group of hospitalized or non-hospitalized persons (very often, in fact members of the United States Navy) is asked to report which of a list of stressful life events they have experienced in the recent past. For each of the events asked about, the investigator has a consensually derived stressfulness weight. These weights were established by pioneer stress researchers, Holmes & Rahe (Holmes & Rahe, 1967; Holmes & Masuda, 1974), who asked numerous large groups to evaluate the amount of change, i.e., stress, caused by each of a series of specified life events in the life of a typical and then averaged all of the judgments. The investigator assigns a total stress score to each of his subjects by noting which events they have checked, multiplying them by their associated consensual weights, and summing their products. Note that the assignment of stress scores has nothing to do with the subjects' own perceptions and evaluations of the events which they have encountered. In the typical study, subjects are also asked what sorts of illnesses they have suffered in the recent past. The investigator assigns a total illness score in a manner similar to his derivation of stress scores, through use of a set of consensually derived weights for a wide variety of commonly recognized diseases, and symptoms (cf. Wyler, Masuda & Holmes, 1970). Finally, the investigator establishes the relationship between stress and illness by observing a significant correlation between the total scores for each. He is also able to bolster his conclusion by noting that hospitalized persons report significantly more recent life stress than non-hospitalized subjects (e.g., Paykel, 1974), and that for those who did become ill, a steady increase in stress levels was observed immediately preceding their illness onset (e.g., Holmes, 1970).

The findings from this area of research have had dramatic impact in both professional journals and popular magazines. In the public media (e.g., Wolfe, 1972), readers are provided with a self-administered and self-scored stress test, and told the likelihood of illness associated with each stress score (e.g., a score above 300 means an 80% chance of getting sick). Readers who have undergone either a severe life change or a series of moderate stresses are warned that if they wish to remain healthy they should cut down on all such stressful encounters with their environment, even to the point of avoiding the drive on the Los Angeles Freeway or putting off raising those touchy issues with their spouses. The concluding message in most of the popular stress and illness literature is consistent and clear: Avoid stress if you want to remain healthy.

For this investigator, this conclusion is both premature and misguided on three fronts. First of all, the advice to avoid stress is inadequate insofar as modern life can be characterized as essentially and necessarily stressful (Toffler, 1970; Brodsky, 1977). A person might indeed be able to keep himself from getting married or taking on a new mortgage, but how can he prevent being affected by current pollution levels, or his being assigned to different jobs when his company signs a new contract? Further, increasing levels of stress bring with them increasing opportunities and potential resources (Kobasa, Hilker, & Maddi, 1977). By avoiding stress, modern persons may also be turning away from chances to better their lives.

My second concern about the conclusion that stress is always harmful has to do with the state of the art in stress and illness research. Although correlations are said to range from .20 to .78, the majority of reports fall below .30, and are often accompanied by standard deviations from four to eight times the size of the mean (Rabkin & Struening, 1976). These data may actually be pointing to exceptions to the rule that illness always follows stress, namely, subjects who score highly on stress without getting sick and subjects who have fallen ill without prior encounter of stressful events. Several investigator-critics in the stress field have also suggested that correlations between stress and illness might be radically strengthened or weakened, if predisposing and mediating factors were taken into account. These might include basic constitutional factors (Hinkle, 1974), social resources (Myers, Lindenthal, & Pepper, 1974), and psychological factors affecting the perception of stress and coping with it (Hinkle, 1974). Not enough attention has yet been paid to the function and interaction of these various factors in the response to stressful life events.

My third complaint with the popular literature's pessimistic approach to stress is based on my own empirical work on the relationship between stressful life events and illness. In a recent study (Kobasa, Hilker, & Madde, 1977) involving all of the middle and upper level executives of a large, public utility, a group of persons were identified who had undergone significant life stress without falling ill. These executives were found to differ significantly from their comparison group, executives who had encountered comparable high levels of stress followed by illness onsets, primarily in terms of personality. The highly stressed but healthy executives were distinctive in basic attitudes, goals, and motivations which were observed to influence the way they perceived, evaluated, and coped with stress.

The study reported in this talk broadens the notion of an interaction among stressful life events, personality, and health status to include the behavior of another professional group. In this study, the correlation between stress and illness is observed in a group of United States Army Officers. It is hypothesized that the relationship between stress and illness is weaker for those officers possessing those personality characteristics identified with the highly stressed but healthy executives, and stronger for those officers who do not share these personality factors.

The specific personality characteristics hypothesized as distinguishing persons who remain healthy under stress from those who fall ill are drawn from existential personality theory (Kobasa & Maddi, 1977) and from the theoretical and research literature on optimal adulthood functioning (Neugarten, 1968; Henry, 1968), coping mechanisms (Lazarus, 1966; Hamburg & Adams, 1974), and psychosomatic medicine (Engle, 1968; Schamle, 1972; Antonovsky, 1974; Hinkle, 1974).

Hypothesis one: Among officers under stress, those who have a sense of commitment to the various areas of their lives will remain healthier than those who are alienated. Committed persons possess the sense of purpose and active involvement with others which minimizes the threat of stressful events, and provides a basis for continual grappling with problems and setbacks (Lazarus, 1966; Antonovsky, 1974; Kobasa & Maddi, 1977). In contrast, the alienated person considers the world worthless and finds insufficient value in stressful events to justify any effort to cope with them (Moss, 1973; Schmale & Iker, 1971).

Hypothesis two: Among officers under stress, those who have a greater sense of control over their lives will remain healthier than those who feel powerless in the face of external forces. Because hardy persons locate the power over occurrences within themselves, they are not overwhelmed when their environment begins to change. The highly stressed but healthy officer is hypothesized to have (a) decisional control, or the capability of autonomously choosing among various courses of action, (b) cognitive control, or the ability to interpret, appraise, and incorporate stressful events into an ongoing life plan, and (c) behavioral control, or a greater number of suitable responses to stress developed through a characteristic motivation to achieve across all situations. In contrast, the highly stressed and sick officer is powerless, nihilistic, and low in motivation for achievement (cf. Averill, 1973).

Hypothesis three: Among officers under stress, those who seek challenge and novelty will remain healthier than those invested in familiarity and security. Those who seek change have well explored their environment, know where to turn for resources in coping with stress, and are well practiced at self-imposed readjustment. Consequently, they can enjoy stress, be less debilitated by it, and recover from it more quickly (Antonovsky, 1974; Maddi, 1975; Moss, 1973). In contrast, those who seek security and familiarity will dislike stressful life events intensely, and be relatively unable to cope with them.

METHODS

Subjects

Twenty-five Army officers (three majors and twenty-two captains), enrolled in a special summer program at a large midwestern university in preparation for ROTC assignments, served as subjects for this stress, personality, and health study. In demographics, the group was notably homogeneous. The typical Army officer subject is easily described as married, in his thirties, father of one or two children, Protestant, but attending church rarely, college graduate with some graduate level training, husband to a woman who works within the home, and member of the infantry branch with six to fifteen years of service completed. Only one of the subjects was black, the other twenty-four were white males. Some diversity was noted in reports of father's occupation. The most frequently reported categories were skilled tradesman ($n=7$) and retired army officer ($n=6$).

Measurement of Stress and Illness

All officers filled out standardized questionnaire measures of stressful life events (Holmes & Rahe, 1967) and symptomatology (Wyler, Masuda, & Holmes, 1970). These are the measures which have provided a good deal of the evidence from which the conclusion has developed that stress increases the likllihood of illness. Each measure includes a wide range of stress events and symptoms, and requires the subject to indicate the frequency and month of their occurrence over the previous three years.

Stress and illness responses were scored by using the consensual weights provided in the Social Readjustment Rating Scale (Holmes & Rahe, 1957) and the weights for illness in the Seriousness of Illness Rating Scale (Wyler, Masuda, & Holmes, 1968).

Measurement of Personality Differences

Approximately one week after their completion of the stress and illness questionnaire, the officers were administered standardized questionnaire measures of many personality dispositions, including those relevant to the hypotheses stated earlier. The investigator's conceptualization of their conditioning effect upon the stress-illness relationship was not shared with the subjects until after questionnaires were filled out and scored.

The tests included the Alienation Test (Maddi, Kobasa &Hoover, 1978), the Internal vs External Locus of Control Scale (Rotter, Seeman, & Liverant, 1962), and a number of scales from the Personality Research Form (Jackson, 1974) and the California Life Goals Evaluation Schedules (Hahn, 1966).

Data Analysis

In determining the impact of personality characteristics on the reaction to stressful life events, the group of officers was split into two at the median on each of the nine personality dimensions conceptualized as relevant to the hypotheses. The correlation between stress and illness for the subgroup above the median personality score (e.g., those high on alienation) was then compared with the stress-illness correlation for the subgroup below the median (e.g., those officers low on alienation). The difference between the high personality and low personality correlation was calculated for each of the nine personality characteristics. A significant difference in correlation was taken to be indicative of personality's significance as a conditioner of the effects of stressful life events on illness onset.

RESULTS

Stress and Illness Scores

The mean stress score for the entire group of Army officers was 982.5, with a standard deviation of 412, a range from 202 to 1802, and a median of 951. The mean illness score was 586, with a standard deviation of 531.8, a range from 0 to 1733, and a median of 562. Comparing officers' average scores with norms available in the literature indicates that every year they encounter sufficient change and demands for readjustment to constitute *major life crisis*. Their illness levels, however, should be characterized as relatively low. A comparison of officers' scores with those obtained from middle and high level executives provides some striking differences. Officers' stress scores are significantly higher ($p<.025$) than the executives' mean stress score of 399, and their average illness score is significantly lower ($p<.025$) than the average executive score of 913.

The three stressful life events most frequently reported by the officers were change in residence (26 reports of), transfer to a new job (22 reports of), and gaining of a new supervisor (21 reports of). The last two events were also found among the executives' most frequently

reported stresses. Among those events reported sufficiently frequently (at least 15 reports of), the three most stressful, i.e., those given the highest consensual stressfulness weights, were marital separation, sexual difficulties, and change in financial circumstances.

The Army group's correlation between stress and illness was remarkably high, with a Pearson product moment figure of .62 ($p<.01$). This expresses a relationship between stressful life events and illness onset much stronger than that reported in most of the stress research, and much stronger than that obtained in the testing of executives, where the correlation was .24 ($p<.025$). Although the officers reported very little illness, the .62 figure indicates that when they do report illness, even a very low level of it, they also report having experienced some kind of stressful life event.

Testing of Personality Hypotheses

Having established that there is a highly significant relationship between the occurrence of stressful life events and the onset of illness for the Army officers, we can now look at whether this relationship is weaker for those officers possessing the personality characteristics associated with health under stress, and stronger for those officers without these characteristics.

Table I shows that personality is a significant conditioner of the effects of stressful life events.

TABLE I
Stress and Illness Correlations for Officers Scoring Above the Median on Personality vs. Those Below the Median

Personality Variable	r for Officers above Median	r for Officers below Median
Control		
External Locus of Control**	.71	.45
Dominance	.65	.66
Leadership	.50	.67
Achievement**	.37	.72
Change		
Interesting Experiences	.64	.51
Cognitive Structure**	.78	.44
Endurance	.43	.42
Security**	.84	.48
Commitment		
Alienation**	.78	.39
–Self	.57	.69
–Work	.50	.68
–Social Institutions	.59	.52
–Family**	.70	.39
–Interpersonal Relations	.58	.49
–Powerlessness	.74	.58
–Vegetativeness	.64	.60
–Nihilism**	.78	.46
–Adventurousness	.61	.60

**Correlation above the median is significantly different from correlation below the median (p .025)

There is support in these results for all three hypotheses stated earlier. Confirming the first hypothesis, officers scoring low on alienation, i.e., officers who are committed, show a significantly lower correlation between stress and illness than those who are high on alienation. The specific form of alienation which characterize those officers who have fallen sick under stress are feelings of estrangement from one's family, and an attitude of nihilism. Those officers who feel involved with their families and who have a sense of meaningfulness less frequently suffer illness following stress. Confirming the second hypothesis are two personality findings: those officers who score low on external locus of control, i.e., those who have a sense of personal control over their lives, and those officers who have a high motivation for achievement show a significantly lower correlation between stress and illness than those who feel that power lies in external forces and who are without achievement motivation. Confirming the third hypothesis, officers low in orientation towards security and in need for cognitive structure show a weaker association between stress and illness than those who desire security and who are cognitively rigid. These variables signify the importance of seeing change as an opportunity and of being cognitively flexible.

These personality findings directly parallel those obtained in the executive study.

DISCUSSION

The implications of these results for persons and institutions interested in curbing the illness provoking effects of stress are clear. Efforts should be devoted to developing attitudes of commitment and meaningfulness, belief that one can control and transfrom the events of one's experience, and perception of change as opportunity and challenge rather than threat.

Certainly, social science has a lot more work to do in the further explication of the interaction of stress, personality, and health status, and in the application of its findings in practical intervention. Other social institutions will hopefully also contribute to this endeavor to illuminate specific ways of fostering the attitudes, goals, and motivations which facilitate both a productive and healthy life.

REFERENCES

Antonovsky, A.,Conceptual and methodological problems in resources and stressful life events. In B.S. Dohrenwend & B.P. Dohrenwend (Eds.), *Stressful life events: Their nature and effects*. New York: John Wiley & Sons, 1974,

Averill, J.R., Personal control over adversive stimuli and its relationship to stress. *Psychological Bulletin*, 1973, *80*, 286-303.

Coelho, G.V., Hamberg, D.A., & Adams, J.E. (Eds.). *Coping and adaptation*. New York: Basic Books, 1974.

Dohrenwend, B.S., & Dohrenwend, B.P. (Eds.). *Stressful life events: Their nature and effects*. New York: John Wiley & Sons, 1974.

Engel, G., A life setting conductive to illness--the giving-up--given-in complex. *Bulletin of Menninger Clinic*, 1968, *32*, 355-365.

Gunderson, E., & Rahe, R. (Eds.). *Life stress and illness*. Springfield, Ill.: Thomas, 1974.

Hahn, M. E., *California Life Goals Evaluation Schedules*. Western Psychological Services, 1966.

Henry, W.E., Personality change in middle and old age. In E. P. Norbeck, D. Price-Williams, & W. M. McCord (Eds.) *The study of personality: An interdisciplinary appraisal*. Chicago: Holt, Rinehart, & Winston, 1968.

Hinkle, L. E., Jr., The effect of exposure to culture change, social change, and changes in interpersonal relationships on health. In B. S. Dohrenwend & B. P. Dohrenwend (Eds.), *Stressful life events: Their nature and effects*. New York: John Wiley & Sons, 1974.

Holmes, T. H., and Masuda, M., Life change and illness susceptibility. In B. S. Dohrenwend & B. P. Dohrenwend (Eds.), *Stressful life events: Their nature and effects*. New York: John Wiley & Sons, 1974.

Holmes, T. H. & Rahe, R. H., The Social Readjustment Rating Scale. *Journal of Psychosomatic Research*, 1967, *11*, 213-218.

Holmes, T. S., *Adaptive behavior end health change*. Medical Thesis, University of Washington, 1970.

Jackson, D. N., *Personality research form manual*. Goshen, N.Y.: Research Phychologists Press, Inc., 1974.

Kobasa, S.C. Hilker, R. R. J., & Maddi, S. R., Remaining healthy in the encounter with stress. *Preceedings of the 37th American Medical Association Congress on Occupational Health*, 1977.

Kobasa, S. C. & Maddi, S. R., Existential Personality Theory. In R. Corsini (Ed.). *Current Personality Theories*. Itasca, Il: F. E. Peacock Publishers, 1977.

Lazarus, R. S., Psychological stress and the coping process. New York: McGraw-Hill, 1966.

Maddi, S. R., The strenuousness of the creative life. In I. A. Taylor & J. W. Getzels (Eds.), *Perspectives in Creativity*. Chicago: Aldine, 1975.

Maddi, S. R., Kobasa S. C., & Hoover, M., *The Alienation test: A structured measure of a multidimensional subjective state*. Manuscript submitted for publication, New York: John Wiley & Sons, 1973.

Myers, J. K., Lindenthal, J. J., & Pepper, M. P., Social class, life events, and psychiatric symptoms: A longitudinal study. In B. S. Dohrenwend & B. P. Dohrenwend (Eds.), *Stressful life events: Their nature and effects*. New York: John Wiley & Sons, 1974.

Neugarten, B., Adult personality: Toward a psychology of the life cycle. *Middle age and aging*. Chicago: University of Chicago Press, 1968.

Paykel, E. S., Recent life events and clinical depression. In E. D. Gunderson & R. H. Rahe (Eds.), *Life stress and illness*. Springfield, Il: Charles C. Thomas, 1974.

Rabkin, J. G. and Struening, E. L., Life events, stress and illness. *Science*, 1976, *194*, 1013-1020.

Rotter, J. B., Seeman, M. and Liverant, S., Internal vs external locus of control reinforcement: A major variable in behavior theory. In M. F. Washburne (Ed.), *Decisions, values, and groups*. London: Pergamon, 1962.

Schmale, A. H., Giving up as a final common pathway to changes in health. *Advances in Psychosomatic Medicine*, 1972, 8, 20-40.

Schmale, A. H. & Iker, H., Hopelessness as a predictor of cervical cancer. *Social Science and Medicine*, 1971, *5*, 95-100.

Toffler, A., *Future Shock*. New York: Random House, 1970.

Wolfe, S. W., Avoid sickness--how life changes affect your health. *Family Circle*, May 1972, pp. 30; 166-170.

Wyler, A. R., Masuda, M., & Holmes, T. H., Seriousness of illness rating scale. *Journal of Psychosomatic Research*, 1968, *11*, 363-375.

Wyler, A. R., Masuda, M. & Holmes, T. H., Seriousness of illness rating scale: Reproducibility. *Journal of Psychosomatic Research*, 1970, *14*, 59-64.

ARMY PSYCHOLOGY IN THE 1970's

E.R. Worthington
Brooke Army Medical Center
Fort Sam Houston, Texas

ABSTRACT

This study examined the impact recent changes in the US Army have had on AMEDD (Army Medical Department) Psychology. In the 1970's U.S. involvement in Vietnam ended and the draft terminated. The size of the Army has shrunk and human resource development programs have proliferated. Prior to the 1970's most AMEDD psychologists entered active duty without previous service and left upon completing their obligation. In the late 1960's veterans were recruited into AMEDD psychology programs; this has proved successful. Clinical psychology has grown 22% while Army strength has declined, the number of field grade psychologists has tripled. In times of decreasing personnel, AMEDD Psychology is increasing and becoming more involved in the Army's human resource programs.

In 1972 United States involvement in the Vietnam conflict, the longest war in our history, was terminated. This complex struggle in Southeast Asia had a profound effect on our society as well as the U.S. Army. The size of the active Army has decreased drastically in the 1970's. Job specialties open to women have increased and the U.S. Military Academy now has female cadets. Reduced promotion, passovers, RIFs and other events associated with the process of turning a large wartime military force into a smaller peace time Army have occurred. The Army has also become an "all-volunteer" force, the draft has ended. It must now compete with the private sector for recruiting and maintaining human resources. This has created an added awareness on the importance of identifying and understanding relative factors of job satisfaction. The supply of human resources is primarily dependent upon two variables; the civilian job market and the Army's ability to meet the needs of its soldiers. These factors have motivated the Army to spend considerable time and energy implementing many human resource development programs. Their stated goals reflect ideas of enhancing human potential, creating job satisfaction and enrichment, increasing organizational effectiveness, and generally improving the quality of life for those in uniform. All of this has transpired in less than six years.

As expected, these events have not passed unnoticed in the Army Medical Department (AMEDD), especially as it applies to AMEDD Psychology. Until 1974 the primary means of procuring Ph.D psychologists was through an AMEDD Graduate Student Program. Each year 5-10 psychology graduate students were recruited into this program. They had to have completed at least one year of graduate school yet be able to complete both their academic requirements and their internship in no more than 3 years. If they met these requirements and were also eligible for commissioning, they could be selected and brought on active duty (as a graduate student) as a 1LT or a Captain. This program terminated in 1974. AMEDD personnel stated the Army was able to find sufficient fully qualified civilian psychologists requesting active duty assignments. Therefore, there was no need to continue to fund a training program. While the program was an unqualified success in procuring new psychologists, the individuals seldom remained on active duty beyond their initial committment. A new training procurement program was implemented in 1978. The three AMEDD American Psychological Association approved internship training sites (Walter Reed Army Medical Center, William Beaumont Army Medical Center and the Army Hospital at Fort Ord, California) will receive six Army interns commissioned from graduate school. Next year (1979) the number of interns is expected to increase to eleven and there is the possibility of developing a fourth intern site.

Basically there are two types of AMEDD psychologists: (1) clinicians--clinical, counseling, educational, etc., psychologists; and (2) researchers--research, industrial, experimental, etc., psychologists. In 1978 the U.S. Army has 96 clinical psychologists and 37 researchers on active duty. This represents almost a 3 to 1 ratio. In 1972 there were 77 clinicians and 38 researchers; a 2 to 1 ratio. These figures indicate the increase in positions for clinicians and the decrease in positions for the researchers. It also shows a 16% increase in total AMEDD psychologists over the past 5 years. During this period the total MSC strength has remained steady with 1850 officers.

Psychology Officer Strength

1972	Clinical	77
	Research	38
	Total	115
1974	Clinical	86
	Research	38
	Total	124
1977	Clinical	94
	Research	32
	Total	126
1978 (1 Feb)	Clinical	96
	Research	37
	Total	133

An examination of the rank structure of AMEDD psychologists reveals almost a 300% increase in field grade strength in the past five years. In the late 1960's and early 1970's the previous Army Surgeon General's Psychology Consultant recognized a need to bring ex-military psychologists into the Army. He felt that if these people could be successfully recruited, their potential for remaining on active duty would be extremely high. This proved true. Several psychology graduate students brought into the program were reserve officers (senior captains and majors) with previous active duty tours in other branches of the Army. These men have remained on active duty and received subsequent promotions. The following chart presents a breakdown by comparing the number of officers on active duty in 1972 to 1977.

RANK	1972	1977
COL	2	1
LTC	3	7
MAJ	3	13
CPT	105	105
LT	2	0
Total Officers	115	126

Another interesting area of comparison is where AMEDD psychologists were assigned in 1972 and where they are being assigned today. Many positions today were not available five years ago. For example, each Army combat division in the U.S. has an AMEDD psychologist assigned to the Mental Health Section of the Medical Battalion. These positions did not exist in 1972. As Vietnam ended, considerable combat strength was shifted from Southeast Asia to Europe. This has assisted in the buildup of mental health professionals in Army units and support organizations throughout Germany. With more psychologists remaining on active duty there are more opportunities to participate in career education programs (i.e., post-doctoral fellowships, the Officers Advanced Branch Course, Command and General Staff College). There has also been an increased demand for AMEDD psychologists to be assigned to non-AMEDD positions. These assignments require the expertise of a behavioral scientist in areas of research design, human resource development, teaching, health care studies, organizational behavior consultation, for example. The following chart depicts various areas of assignments reflecting changes over the past five years.

Assignment	Number of personnel assigned	
	1972	1977
Europe	7	18
FORSCOM	0	10
Career Education	0	3
Non-AMEDD positions	6	10
No Change		
Health Services Command		
Academy of Health Sciences		
Decreased Emphasis		
Pacific	5 (1974)	0
Research positions	37	28
Graduate Student Program	55	10

This has been a brief review of changes in AMEDD Psychology during the last five years. During this time period the Army has decreased in size; the AMEDD strength has remained stable with about 1850 officers assigned; and AMEDD Psychology has increased in size (115 to 133). This suggests that there exists a recognized need for Army behavioral scientists. Today AMEDD psychologists are serving in a variety of positions throughout the world. Current available assignment locations are:

1978 Assignment Locations

Clinical Psychologists (68S)

Health Services Command (CONUS)	63
U. S. Military Academy	2
TRADOC (CONUS)	4
FORSCOM (CONUS)	11
Europe	16
Korea	1
Total	97

Research Psychologists (68T)

Development & Research Command	12
DOD/Joint	1
Office of the Surgeon General	23
TRADOC	3
Total	39

One problem is beginning to develop. During the 1960's and early 1970's there were few psychology officers. Most assignments were established for the rank of captain with some positions for majors. Today's majors and lieutenant colonels do not find these positions challenging after two or three years. These officers are now searching for new areas of responsibility, which in most cases, requires leaving the immediate area of psychology. One colonel has been assigned to the faculty of the Army War College; a lieutenant colonel was selected for appointment as Chairman of the Leadership Department at the USMA (this included a promotion to colonel). Another officer, a major, was selected to attend the USA Organizational Effectiveness Staff Officer course at Fort Ord, California. Upon completion of the 15 week school in organizational consultation, he will be assigned to Tripler Army Medical Center as the internal organizational development consultant. These are but some of the possibilities available in the late 1970's for AMEDD psychologists.

WHY DO SOME ARMY PSYCHOLOGISTS LEAVE THE SERVICE?

A. David Mangelsdorff, Ph.D.
Psychology Consultant (MobDes)
US Army Health Services Command[1]
Fort Sam Houston, Texas

ABSTRACT

Active Duty Army Psychologists (114 of 130) and psychologists who had left the service since July 1974 (48 of 69) were surveyed on their perceptions of the Army. Responses were input into discriminant analyses to determine which issues significantly discriminated between the Active Duty and Left Service groups. The discriminant function developed correctly classified individuals 84.0% overall into their respective groups.

The retention of health care professionals like psychologists and physicians has been a continuing concern for the military services. The problem of retaining military medical officers has been addressed at varying times (Hedlund, 1968; Cooke and Mixson, 1967, 1971; Baker, 1969; Boyson, 1967; Winkler, 1968; Braustein, 1974; Mangelsdorff and Hubbart, 1976; Watson, 1976; Jorlett, 1975; Lanier, 1975; Krause, 1978; Dully, 1974). Among the recurrent findings influencing physicians to leave the service were: inadequate pay, possibility of command or administrative assignments, lack of amount of participation in making decisions affecting own career, lack of sense of belonging to the community and social life of the military, and poor facilities.

The purpose of this study was to document what factors induce some psychologists to remain in the Army and what causes other psychologists to leave the service. The intents are to determine what factors would increase the likelihood of psychologists remaining on active duty and to document what degree of job satisfaction Army psychologists report.

METHOD

Subjects. Active duty Army psychologists (N = 130) and psychologists who had left the Army since July 1974 to November 1976 (N = 69) were included.

Procedure. Each psychologist was mailed a survey instrument which requested demographic background, military background, attitudes toward military career, levels of satisfaction, retention factors, and the Job Descriptive Index (Smith et al, 1969). A personal letter requesting participation was provided in addition to an official request from the Academy of Health Sciences. Return self-addressed envelopes were provided. Active duty psychologists were asked to use a 7-point Likert scale to rate how satisfied they felt about a variety of issues. Psychologists who had left the Army between July, 1974 and November, 1976 were instructed to answer the survey questions as though they had six months remaining in their active duty obligation. A follow-up was conducted requesting psychologists who probably had not participated to indicate whether they would respond, needed a questionnaire, or would not return the survey.

RESULTS

Sample Characteristics. Of the active duty psychologists, 114 of 130 (88%) responded; of the psychologists who left the Army, 48 of 69 responded (70%) returned the questionnaire. The average age of the sample was 31 years. See Table 1.

Job Descriptive Index. The Job Descriptive Index scales (JDI) were scored and analyzed as dependent variables using analysis of variance (ANOVA) procedures. The JDI consists of five subscales measuring different aspects of an individual's attitudes toward their job. The subscales include: WORK, SUPERVISOR, CO-WORKERS, PAY , and PROMOTIONS. There was only one significant difference between Active Duty and Left Service groups on the JDI scores, that for PROMOTIONS, where the Active Duty group reported significantly greater satisfaction than the Left Service group ($F = 4.588$ (1 153), $p = .034$).

Discriminant Analyses. Stepwise (Wilks procedure) discriminant analyses were conducted to separate Active Duty and Left Service groups using the Satisfaction set attitude items (7-point Likert), the five subscales of the Job Descriptive Index, and the Demographic section items. Items having F values equaling or exceeding 1.00 werre extracted as discriminator items. Also items selected had to have been responded to by at least 145 psychologists. The discriminant function developed correctly classified individuals 84.0% overall into their respective groups. The Active Duty group (N = 114) was correctly classified 85.1%, while the Left Service group (N = 48) was correctly categorized 81.3%. The discriminant function developed was significant ($X^2 = 74.69$, $p < .01$).

DISCUSSION

Table 2 extracts the most significant discriminants as entered in the step-wise procedure. The most significant discriminant was satisfaction with *Personal control over how my own career develops,* followed by the extent of *Participation in military oriented social activities.* With respect to the perception of personal control, Nord (1977), Seeman (1972), and Form (1975) suggest that the power individuals exercise over the situations that affect them, relative to their expectations of how much influence they should have, may be an important determinant in job dissatisfaction. Since all of the psychologists who left the service were captains (the one exception retired from the Army), there may have been some feelings of powerlessness in their position by some of the psychologists who left the Army. Shephard and Panko (1974) note that power-deficit workers had less commitment to organizational goals. Of the Left Service group, all but one rank ordered themselves as psychologists first, and all placed their rank ordering of self as military officer as lowest.

With competition from civilian jobs perhaps offering higher pay, independence, stability, opportunities for self-improvement, or some facet not otherwise found in the military, an Army psychologist may choose not to remain in the Army. In addition, factors supporting a negative view of the military may influence the decision to leave the service. The specific issues discernible from the analyses (using the standardized discriminant function coefficients) which suggested an individual would be in the Left Service group were the degree of satisfaction with: *The amount of leisure time I have available, Having rank commensurate with experience, Availability of civilian non-federal job opportunities, Opportunity for self-improvement outside my job,* and *Personal accomplishments as a military psychologist.* In addition, the extent *Feel a sense of belonging to the community and social life of the military* and probability of *Entering military service for military pay and fringe benefits relative to those available as a civilian at that stage of my professional education* were significant contributers of Left Service psychologists. The Left Service psychologists generally did not feel part of the military. Social and professional support for the decision (whether to leave or remain in the service) is critical. It would be quite difficult to remain in the service without support from a spouse (or significant other) and or co-workers.

In addition, job satisfaction, accomplishments, and personal and or professional growth must be considered. Issues which contributed to categorizing a psychologist in the Active Duty group dealt with the development of an identity as a career military officer concerned with career progression, how the current assignment allowed for personal accomplishments as a military officer, and how the individual viewed himself. Specifically, *Likelihood promoted* was of great concern to the individual (as evidenced by the large discriminant function coefficient). Further support for this concern with promotions comes from the analysis of variance between the Left Service versus the Active Duty group scores on the JDI PROMOTIONS subscale, where there was a significant difference between the groups ($p < .05$). Participation in military oriented activities may be perceived as possibly affecting one's military career. The benefits of joining the Army for educational opportunities must also be emphasized. A sizeable number of psychologists would consider remaining on active duty to take advantage of them. For Army physicians, the availability of residencies and post-residency specialized fellowships has been cited as a very significant factor in why some physicians remain in the Army (Krause, 1978; Whelan, 1974).

FOOTNOTE

[1]Health Care Studies Division, Academy of Health Sciences, Fort Sam Houston, Texas 78234 is the work address.

REFERENCES

Available on request.

TABLE 1
DEMOGRAPHIC CHARACTERISTICS OF ACTIVE DUTY (AD), LEFT SERVICE (LS), EXPERIMENTAL, AND CLINICAL GROUPS

	EXPERIMENTAL		CLINICAL	
SEX	**AD**	**LS**	**AD**	**LS**
Male	38	13	75	34
Female	0	0	1	1
RANK	**AD**	**LS**	**AD**	**LS**
1LT	0	0	1	0
CPT	32	13	60	34
MAJ	5	0	12	0
LTC	1	0	2	0
COL	0	0	1	1
SOURCE OF COMMISSION	**AD**	**LS**	**AD**	**LS**
ROTC	19	8	34	14
GSP	9	5	22	16
OCS	0	0	3	0
SVC ACAD	0	0	6	0
DIR COMM	5	0	4	4
ROTC & GSP	2	0	7	0
DC & GSP	2	0	0	1
YRS PRIOR ACT MIL SVC	**AD**	**LS**	**AD**	**LS**
None	25	11	47	33
1-4 YRS	11	2	14	2
5-9 YRS	2	0	6	0
10 & MORE	0	0	8	0
DOCTORATE CONFERRED	**AD**	**LS**	**AD**	**LS**
NO	1	0	11	0
YES	37	13	65	35

TABLE 2

Significant Contributers To Discriminant Function Developed Stepwise (Wilkes Procedure) To Discriminate Active Duty (AD) Versus Left Service (LS) Groups

Step	Item Content of Variable Entered/ Removed	F to Enter or Remove	Standardized Discriminate Function Coefficient	AD Mean	LS Mean	Uni-variate F Ratio
1	Personal control over how my own career develops	21.97	-.328	3.84	2.17	21.97
2	Participation in military oriented social activities	11.74	-.424	3.13	1.92	13.44
3	The amount of leisure time I have available	6.04	.456	4.47	5.22	5.58
4	Enter military service for educational benefits after service	5.66	-.451	2.51	1.57	8.35
5	Availability of civilian non-Federal job opportunities	4.67	.343	4.34	4.75	1.26
6	Rank ordering of self as a military officer[a]	4.21	.189	2.78	3.00	7.10
7	Extent feel being utilized professionally	3.79	-.246	4.94	3.75	9.93
8	Having opportunity for travel	3.10	-.309	4.88	4.05	6.98
9	Months in present assignment[b]	2.78	.214	21.57	29.95	6.61
10	Personal accomplishments as a military psycholigist	1.41	.229	4.86	4.32	2.69
11	Likelihood promoted	1.10	-.458	5.03	3.77	9.95
12	Extent of time spent on special military duties	1.09	-.363	3.51	1.85	19.48
13	Belonging to community and social life of military	1.52	.236	3.19	2.12	9.69
14	Having rank commensurate with experience	1.28	.373	2.57	2.02	2.37
15	My co-workers positive attitudes toward military	1.28	-.282	4.51	3.92	3.66
16	Enter military for military pay and fringe benefits available[c]	1.31	.298	3.25	2.62	2.58
17	Enter military for graduate student program[c]	1.21	-.344	4.38	3.55	2.74
18	Having stability in my current assignment	0.93	-.251	4.94	3.77	8.40
19	Opportunity for self-improvement outside my job	1.36	.256	4.59	4.15	2.08
20	My positive feelings toward my supervisor	0.61	.201	4.85	4.40	1.28

NOTES:
1. Postive standardized discriminant function coefficients categorize individuals into Left Service group.
2. Unless otherwise specified by superscripts, a seven point Likert scale from minimum (1) to maximum (7) was used.
   a. Rank ordering could be 1, 2, or 3.
   b. A numerical entry.
   c. Used a 7 point Likert scale from Low (1) to High (7).

A JUDGMENT ANALYSIS APPROACH TO EXAMING THE "WHOLE PERSON CONCEPT"

Adrian M. Harrell and Phillip Glenn
Air Force Institute of Technology
Wright-Patterson AFB, Ohio

ABSTRACT

This paper reports the results of an experiment that investigated how the "whole person concept" influences the promotability judgments of junior and senior groups of Air Force officers. The subjects, who were randomly selected from officers attending the Squadron Officer School and Air War College, completed a decision making exercise that involved using multiple criteria to judge the promotability of a group of hypothetical captains. The research approach was based upon the Brunswik Lens Model. A full factorial design was used. It was concluded that (1) Air Force officers incorporate the "whole person concept" into their promotability judgments, (2) junior officers weight the promotability factors differently than senior officers and (3) little agreement exists within officer groups as to how the promotability factors should be weighted.

The basic authority for promoting Air Force officers is found in the United States Code, Title 10 (USC 10). This statute allows the Secretary of the Air Force to establish a promotion board which recommends officers for promotion. The "best qualified." The importance to be associated with each of the individual criteria is apparently left to the discretion of the individual board members.

TABLE 1
Whole Person Concept (Source: HQ AFMPC)

Total Evaluation

Factors	Evaluate
Performance	OERs
Education	Level/Utilization
Breadth of Experience	Where/When/What
Job Responsibility	Scope/Exposure
Professional Competence	Expertise of Specialist
Combat/Achievement	Awards/Decorations
Leadership	Staff Command

Other Factors/Evaluations May Apply

Problem Statement

It seems reasonable to expect that the behavior of many Air Force officers is motivated by the desire to acquire the attributes they believe will influence their promotability. With this in mind, this study examined how selected Air Force officers were influenced by decision criteria derived from the "whole person concept" in judging the promotability of a group of hypothetical captains to the rank of major. Three research hypotheses were considered.

1. Officers do include the "whole person concept" into their promotability judgments by using all the promotion criteria in reaching their decisions.
2. Junior officers place the same relative weights upon the criteria used for their promotability judgments as do senior officers.
3. Individual officers place the same weights upon the criteria used for their promotability judgments, resulting in homogenous judgment policies within officer groups.

These three hypotheses were investigated by analyzing the promotability judgments reached by officers attending the Squadron Officer School and the Air War College.

Research Methodology

The research approach used in this study is conceptually based upon the Brunswik Lens Model (Brunswik, 1952). This approach for analyzing human judgment has been described in detail by Dudycha and Naylor (1966), Beach (1967), and Hoffman, Slovic and Rorer

(1968). Slovic and Lichtenstein (1971) have presented a review of the pertinent literature. Using this approach meant asking officers attending the Squadron Officer School and the Air War College to complete a decision making exercise. This exercise involved judging the promotability of a number of hypothetical captains to the rank of major. An example of a typical decision case is presented in Table 2. Some additional information about each of the decision criteria was also presented to each of the subjects.

TABLE 2
Sample Decision Case

Captain #1

Decision Criteria	Individual Performance
PME HISTORY	Squadron Officers' School
ASSIGNMENT HISTORY	Entirely at base level
AERONAUTICAL RATING	Pilot
OER RATINGS	2, 1, 1
FORMAL EDUCATION	Masters degree

DECISION#1

1...........2..........3..........4..........5..........6...........7

Low Promotability	Moderate Promotability	High Promotability	Very High Promotability

It was explained that the Officer Effectiveness Report (OER) ratings should be considered the most recent of a series of three ratings. Individuals received ratings of either 2, 2, 1 or 2, 1, 1. Only 28 percent could receive a rating of 1. As many as 78 percent could receive a rating of 2. It was assumed that the officers whose promotability was being judged all received the same ratings except for those shown. The formal education possessed by each individual was either a bachelors degree or a masters degree. Each individial had either completed the Squadron Officer School or had no professional military education or record. Each individual had either completed a headquarters assignment or else had been assigned entirely at the base level. Each individual was either a pilot or did not possess an aeronautical rating.

It should be pointed out that the hypothetical officers being judged were not intended to be representative of all captains, and the subjects were informed of this. In addition, the Officer Effectiveness Ratings were deliberately designed so that all of the hypothetical captains would have ratings that were quite similar. The individuals being judged differed in this category only by a single rating, the most current rating.

A full factorial design was used. With two measurement levels and five criteria, this meant each subject was presented with 32 decision cases ($2^5 = 32$). With the judgment analysis approach, the relationship between an individual's decisions and the criteria he uses to arrive at those decisions is usually examined with multiple regression analysis, the approach used here. The decision criteria are represented by the independent variables and the individual's decisions (the overall judgment of promotability here) are represented by the dependent variable. Although linearity is not a requirement of the Brunswik Lens Model, a number of empirical studies have established that only a small percentage of the explainable variance can normally be attributed to nonlinear effects (Hammond and Summers, 1965; Hoffman, Slovic and Rorer, 1968). The calculation procedure described by Hoffman (1960) was used to determine the relative weights each subject placed on each of the five criteria in reaching his decisions. The orthogonal predictor vectors resulting from the full factorial design meet the requirements described by Ward (1962) for this approach to be used. Essentially, this procedure involves partitioning the explainable variance (R^2) appropriately to each independent variable. The relative weight associated with a variable is the percent of the explainable variance attributed to that variable.

The weights the two groups of officers placed upon the five criteria were compared by comparing group multiple regression models using the F-Test due to Chow (1960). This test compares the beta coefficients of multiple regression models. Student's t-test was used to compare individual relative weights.

Results of the Analysis

The initial research hypothesis proposed that officers do include the whole person concept into their promotability decisions by using all the promotion criteria in reaching their decision. An F-test of the significance of the regression coefficients include into the computer algorithm revealed that both officer groups placed significant weights ($p < .001$) upon all five decision criteria. (See Table 3). It should be noted, however, that almost 80 percent of the explainable variance was associated with only three of the five variables for each of the two groups.

The second research hypothesis proposed that junior officers place the same relative weights upon the criteria used for their promotability judgments as do senior officers. When group regression models were compared, significant differences in how the criteria were weighted were observed (F-test, $p < .05$). Student t-test comparisons revealed that professional military education, officer effectiveness report ratings and formal education were all weighted differently by the two officer groups (see Table 3).

The third research hypothesis proposed that individual officers place the same relative weights upon the criteria used for their promotability judgments, resulting in homogenous judgment policies within officer groups. The squared multiple correlation coefficient (R^2) can be interpreted to examine this hypothesis. As individuals, the officers attending the Squadron Officer School and the Air War College implemented their particular judgment policies quite consistently, resulting in mean individual R^2 values of .76 and .79 respectively. Relatively low R^2 values (.44 and .40) were observed when these officers were considered as groups, however. This comparison reveals that important differences existed between the judgment strategies employed by individual officers within each of the groups. These data are summarized in Table 3.

TABLE 3
Results of Analysis of Data

Officer Group	Promotion Factor	Relative Weight	Beta Weight Significant (.01 level)	Individual R^2	Group R^2
Squadron Officer School	OER	20*	yes	.76	.44
	Formal Education	23*	yes		
	PME History	35*	yes		
	Assignment History	13	yes		
n = 99	Aeronautical Rating				
Air War C College	OER	41*	yes	.79	.40
	Formal Education	13*	yes		
	PME History	22*	yes		
	Assignment History	15	yes		
n = 85	Aeronautical Rating	09	yes		

*t-test, significantly different at .01 level than corresponding relative weight for the other group.

Conclusions and Implications

Let us examine the conclusions that follow from the results of the analysis and the implications of these conclusions.

Both officer groups incorporated the "whole person concept" into their promotability judgments. This implies that officers seeking to improve their promotability should be expected to plan their careers so as to acquire the attributes they feel are associated with the "whole person concept." Since three of the criteria accounted for almost 80 percent of the explainable variance, it should be expected that officers will concentrate upon officer effectiveness ratings, professional military education and earning a graduate degree to a greater degree than the other factors.

Junior officers differed from the senior officers in how they weighted the decision criteria. The junior officers considered professional military education and formal education more important than did the senior officers. The senior officers considered the officer effectiveness report ratings more important than did the junior officers. When examining these differences, one should keep in mind that in practice it is the senior officers who decide upon the promotability of the junior officers. Consequently, this outcome implies that junior officers overrate the value of attending the Squadron Officer School or earning a masters degree towards improving their promotability.

Important differences existed in how the individual officers within each of the two groups weighted the five criterias. This outcome implies that little agreement exists among either junior officers or senior officers about the importance to be placed upon the various promotion criteria which may be derived from the "whole person concept."

REFERENCES

Beach, Lee Roy, "Multiple regression as a model for human information utilization." Organizational Behavior and Human Performance, 2:276-289 (August 1967).

Brunswick, Egon, *The Conceptual Framework of Psychology*. Chicago: The University of Chicago Press, 1952.

Chow, Gregory C, "Tests of equality between sets of coefficients in two linear regressions." Econometrica, July 1960, 28, 581-605.

Dudycha, Linda W. and James C. Naylor, "Characteristics of the human inference Process in Complex Choice Behavior Situations," *Organization Behavior and Human Performance*, Septenber 1966, 110-128.

Hammond, K.R. and D.A. Summers, "Cognitive dependence on linear and nonlinear cues." *Psychological Review*, 1965, 72, 215-224.

Hoffman, Paul J, "The Paramorphic Representation of Clinical Judgment." *Psychological Bulletin*, March 1960, 57, 116-131.

Hoffman, Paul J., Paul Slovic, and Leonard Rorer, "An analysis of variance model for assessment of cultural cue utilization in clinical judgment." *Psychological Bulletin*, May 1968, 69, 338-349.

Slovic, Paul and Sarah Lichtenstein, "Comparison of bayesian and regression approaches to the study of information processing in judgment." *Organizational Behavior and Human Performance*, 1971, 6, 649-744.

United States Code. Title 10—Armed Forces. Washington, D.C. 1970 edition.

Ward, Joe H., Jr., "Comments on the paramorphic representation of clinical judgment," *Psychological Bulletin*, January 1962, 74-76.

THEORY-BASED PREDICTIONS OF DECISIONS TO RE-ENLIST IN A NATIONAL GUARD UNIT

Thomas W. Milburn and Robert S. Billings
Ohio State University

Like individuals, most task-orineted organizations from time to time face problems severe enough and potentially threatening enough to organizational control, resources, or functioning as to fit the technical label of crisis. Technically, organizational crises precede potential disruption severe enough to be labeled disaster; they call for decisions that, if swiftly and effectively implemented, will forestall disaster; and they involve some degree of surprise so that previous plans are inadequate. From a technical point of view organizational crises are not merely turning points as is the case with a medically sick individual who now begins to mend or die, or developmental choice points as with Erikson's stage points such as trust, autonomy, mastery, identity, or integrity. Nor are they necessarily only episodes of increased tension for their participants, stressful peak experiences for those sufficiently involved in them, conflicts either, on the one hand, to be resolved and problems to be solved successfully as was the case for the U.S. at the time of the Cuban missile crisis, nor, on the other, situations that are merely explosive preludes to periods of disruptive systemic change and high social cost, as was the case with those events covering the six weeks in 1914 following the assassination of Archduke Ferdinand preceding the destruction of World War I. We expect some crises to be chronic ones that do not readily mobilize efforts to resolve them, debilitating in quality rather than exhilarating as so often is the case with acute crises.

We are engaged in a study of organizational crises though the one we shall describe today is a little atypical in being slow and chronic rather than acute. The organization in question is a military one, the National Guard units for one populous state. The situation we research investigators have labeled a crisis concerns the diminishing size of this voluntary organization in a year when forty-five percent of its membership came up for re-enlistment; it is a retention problem. Further reduction in size could portend a reduction in authorized strength and federal support. If the changes occurring are symptomatic of a long-term trend, the survival of the organization as it stands could eventually be at stake. So the situation is one of threat to the organization. Depending upon how long it takes to gather data, analyze them, make and implement appropriate policy steps, there is, or is not, time pressure perceived by participants in the situation. First there must be some systematic exploration of the situation leading to understanding of the problem; then decisions must be made and relevant actions carried through to deal with it.

It has not been altogether clear to us to what extent the situation we are investigating has been one of surprise for the organizational executives. Clearly there has not yet been a plan constructed to deal with the situation; at least the situation is one of uncertainty. To some degree the uncertainty is compounded by a recognition of the need to recruit and retain minorities and women in the organization. So a different organization is called for.

Fortunately, there is a systematic research literature on such topics as military organizanional climates (e.g., Moos, 1975) and on the prediction of performance and retention in the military service (Hoiberg & Pugh, 1977). To understand better the nature of the problems encountered by this organization and potential appeals available to increase retention, all in such a way that we could formulate and test hypotheses in ways that could enable us to begin to define the problems and alternative solutions in policy-relevant terms, we conducted twenty several-hour interviews of members of the organization, starting at the top, at several sites, then constructed, pretested, modified, and administered the first of a two-wave panel sample survey to members of the organization. Although the letter was signed by the General in charge, the return rate was only sixty percent. Still, in absolute terms, our return is substantial. We did oversample for minorities and women. Percent of return was in part a function of unit, some performing quite well, others performing not at all.

It seemed worthwhile to seek to predict who would probably leave the organization, who would most likely stay, and who was undecided. If limited resources are available to deal with the problems, a triage approach would appear to be most appropriate. Some persons are likely to be quite satisfied; it is unnecessary, though perhaps just, further to reward or to attempt to satisfy them. Some persons are so clearly decided upon leaving that to expend effort to persuade and satisfy them would seem a lost cause, a waste of resources. But those who are undecided are another matter.

Hypotheses

Several hypotheses occur to us:

H1 Individuals electing to stay with the organization will report different likes and dislikes than those electing to leave. More specifically,

H2 Some aspects of organizational structure and program that those electing to stay tend to regard as good, those planning to leave will regard as bad. These hypotheses, if borne out, would have implications for selection. Changing selection procedures could contribute to improved retention by seeking persons who are likely to like the organization. Additionally, providing information about the organization's bad as well as good points at the time of the decision to enter it does apparently contribute to retention. That is, those persons who are informed *from the start* concerning an organization's bad points as well as its good ones are more likely to decide to stay.

Another hypothesis of interest would have rather different policy implications:

H3 Individuals electing to stay with the organization will report different experiences (and ones which might be inferred to prove more reinforcing) than will those electing to leave. If supported by our data, this hypothesis is far more optimistic, even for the short run; it would suggest that carefully implemented policy changes, if made swiftly, could change the appeal of the organization rather quickly.

Another hypothesis would have similar and conceivably triage-like implications:

H4 Persons will report differentially satisfying experiences within the organization as a function of which units they are in. If supported, this hypothesis would suggest that efforts to change policy and practice in the organization could most effectively be concentrated upon some units to the exclusion of others. (Of course, the principle of concentration of effort to achieve certain objectives is a familiar and ancient principle of military strategy.)

Results and Discussion

The findings from the initial wave of our survey--and before we discover how accurate are our predictions of leaving or staying based upon stated intentions several months in advance--are quite clear. All of our hypotheses are supported. Of most immediate interest are hypothesis three and four. Specifically, those individuals who have decided to stay report *having* experiences which those intending to leave (tend to) *have missed*, namely conversations with and performance feedback privately from their commanding officers, other commissioned officers, and the noncommissioned officers in their units. Those in the former category express much more satisfaction with their experiences with career counselors than do those in the latter group. Knowing that our respondents report differentially on these matters by unit suggests ways to supply these reinforcing stimuli rather quickly to increase organizational strength.

THE AIR FORCE MEDICAL EVALUATION TEST (AFMET) PROGRAM: A THREE-PHASE ASSESSMENT PROGRAM

Thomas P. O'Hearn
Lackland Air Base, Texas

Introduction

Involuntary separation from the service before completion of the first enlistment has been a long-standing problem for the military. Each service has sought ways to identify those who would fail to function effectively in the military environment (Bucky & Edwards, 1974; LaChar, et. al., 1974; Guinn, et. al., 1975; Plag, 1962). At the present time, with personnel costs running over one half the defense budget, approximately one airman in four does not complete his initial tour of duty. The cost in wasted manpower and training dollars can only be estimated. The problem of attrition starts early in the training process. Approximately eight percent of the incoming airmen do not complete basic military training. The current cost of keeping one airman in basic military training for one day is estimated at $88.00. The thirty training days of basic training average out to about six weeks at Lackland Air Force Base at a cost of $2,640.00. These are training costs and do not include costs of recruiting and travel to San Antonio. The early identification of those trainees who fail to complete basic training would seem, from both an economic and mental health standpoint, to be in the best interest of the Air Force. The Air Force Medical Evaluation Test Program (AFMET) was designed to identify within the first several days of training those airmen who appear to be at risk for completion of basic training and/or effective Air Force service. In 1972, the Department of Mental Health, Wilford Hall USAF Medical Center, Lackland Air Force Base, Texas, addressed itself to the problem of identifying those airmen who, due to psychological, emotional or personality difficulties, would be at risk for completion of basic military training or effective military service.

Method

A 100-item-test named the History Opinion Inventory (HOI) was validated with a sample of 14,000 basic military trainees. The utilization of this instrument enabled the researchers to identify an "at risk" group, forty-five percent of whom, for various reasons, left the Air Force before their initial enlistment (LaChar, et. al., 1974; Guinn, et. al., 1975). In 1975, a research program was devised utilizing the HOI. The program of assessment at that time consisted of three phases. Phase One consisted of administering the History Opinion Inventory to all incoming basic trainees following arrival at Lackland Air Force Base. The test was administered to the trainees before their assignments to the basic training squadrons. Phase two consisted of administering, to those airmen who scored in the top seven percent of HOI, a Minnesota Multiphasic Personality Inventory (MMPI) and a standard psychiatric interview. The latter consisted of a mental status exam and questions designed to assess functioning prior to entry into the Air Force. The interviw data were broken down into thrity-four symptom categories and six historical/behavioral categories. Each category was scored on a five-point scale ranging from normal to severe. The data were recorded on an Optiscan sheet and given the designation of the Standard Report of Interview (SRI). The MMPI and SRI interview data were then reviewed by a clinical psychologist and a determination was made whether an individual would be returned to duty or be evaluated further in Phase Three. Those selected for a Phase Three evaluation were sent to the Mental Hygiene Clinic for further assessment. These evaluations were conducted in the same manner as for those individuals who had been referred to the clinic by squadron personnel or other base agencies. The evaluation usually consisted of written behavioral observations by squadron personnel, a mental status exam, intensive interview, and testing when indicated. Following the Mental Hygiene evaluation, the individual was either returned to duty, returned to duty with follow up by Mental Hygiene, referred for evaluation by Inpatient Psychiatric Service, or referred to a number of other agencies who are tasked with the evaluation of certain problems. For example, if a trainee had a significant drug or alcohol problem, he would be referred to Social Actions for appropriate disposition.

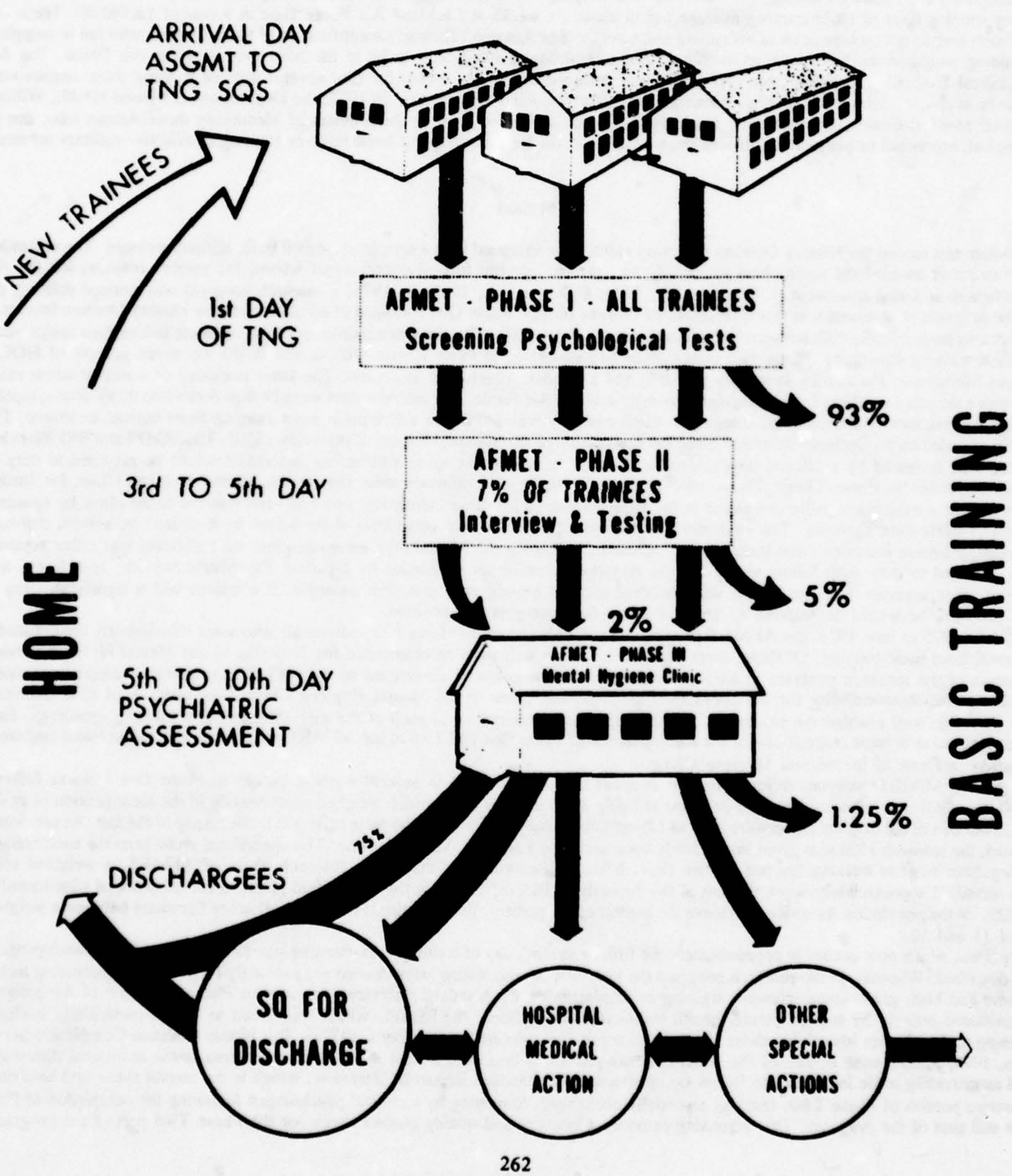
From June 1975 to June 1976, the AFMET program identified four hundred forty-four individuals who were subsequently recommended for removal from basic training. Of these, three hundred six individuals were recommended for discharge by the Mental Hygiene Service. The success of the research program in early identification of low potential airmen led to the AFMET program becoming operational in October 1976. Responsibility for the operational program was given to the Mental Hygiene Clinic. The addition of nine technician slots to the clinic staff enabled the placement of one mental health technician in each of the then thirteen basic training squadrons. Each squadron technician became responsible for the administration of Phase One and Two of the AFMET program as well as the initial evaluation of squadron referrals to the Mental Hygiene Clinic.

The present AFMET program differs from the original research program in several ways. Changes in Phase One were as follows. The 100-item HOI was condensed to a 50-item test utilizing 36 of the 100 items which weighed most heavily in the identification of at-risk airmen. Fourteen of the original items were used as fillers. Other changes in Phase One were reflected in the timing of the test. As previously mentioned, the research HOI was given immediately upon arrival at Lackland Air Force Base. The operational phase tests the basic trainees after they have been in training for one or two days. It is significant to note that in the research phase of AFMET, a weighted score of nine identified approximately seven percent of the population (Bloom, 1976). In the operational phase, a cut-off score of nine identified about 22% of the population. In order to achieve the highest seven percent, the operational phase cut-off score fluctuates between a weighted score of 11 and 12.

Phase Two, which now occurs in approximately the fifth or seventh day of training, has changed significantly from the research program earlier described. Whereas, in the research program the interviewing and testing were carried out primarily by personnel processing technicians who had been given some intensive training in administering the standard interview, the current Phase Two part of the program is administered entirely by trained mental health technicians. In addition, the MMPI, which was found to be not particularly useful as a screening tool in further identifying those airmen who are at risk, was replaced (Bloom, 1977a). The Bloom Sentence Completion Survey (Bloom, 1975) is now used in lieu of the MMPI. It has proven to be useful in that it yields scores across seven attitudinal dimensions as well as providing some leads for the follow-on interview. The Standard Report of Interview, which is the mental status and behavioral observations portion of Phase Two, remains essentially unchanged. Screening by a clinical psychologist following the completion of Phase Two is still part of the program. This screening provides a quantity and quality control check for the Phase Two part of the program.

FIGURE 1.

AFMET-1977



The Phase Three portion of the program is essentially unchanged from the former research program. Individuals who are selected for Phase Three of the AFMET program bring with them to the Mental Hygiene Clinic their medical records and a report from their squadron with comments and observations of their progress and problems, if any, in the basic training environment. The Phase Three part of the program is typically a more global assessment with particular attention to present functioning. Phase Three usually occurs somewhere in the tenth or eleventh day of training and by this time the initial stress of basic training for many trainees is over. In Phase Three, the trainee is given an extensive interview and usually tested with a variety of psychological instruments. Although any staff member may conduct a Phase Three evaluation, it is usually conducted by one of the Officer staff of the Mental Hygiene Clinic which consists, at present, of three social workers and one clinical psychologist. This phase is usually completed in one day. However, evaluations do, at times, take up to several days and may involve a number of assessments by various tests, individuals or agencies. The dispositions from the Mental Hygiene portion of the assessment process are essentially unchanged from the research program. That is, the individual is either returned to duty, referred to another agency for disposition and or treatment, or recommended for discharge by the Mental Hygiene staff.

Results and Discussions

The following are the data generated by the research program in 1975 and the operational program of 1976-1977.

TABLE I
COMPARISON OF DATA FROM RESEARCH AND OPERATIONAL AFMET PROGRAMS
Lackland Air Force Base, Texas

	Research (1 Jun 75-76)	Operational (1 Oct 76-77)
HOI's Given	80,732	73,666
Phase II Selectees	5,367	4,918
Phase III Selectees	1,331	1,054
MHC Recommended Discharges	306	521
Other recommended Discharges	138	336

As can be seen from Table I, both the research and operational programs were successful in the early identification of a substantial number of high-risk trainees. The advantages of the operational program over the research program seem to rest in the amount of individuals referred from Phase Two to Phase Three and the percentage of those individuals who were subsequently recommended for discharge. In the research program, utilization of personnel processing technicians for interviewing in combination with the giving of the MMPI as a screening device led to a substantial number of individuals being referred to the Mental Hygiene Clinic who were subsequently returned to duty. The operational phase of the program which utilized trained mental health technicians and a sentence completion test resulted in fewer referrals to the Mental Hygiene Clinic, but a larger percentage of those referred being recommended for discharge. As can be seen, both the research program and the operational program identified individuals who were subsequently recommended for discharge by other agencies, but who were initially identified by the HOI. The bulk of these discharges appeared to be for medical reasons. Medical problems which existed prior to entry into service, and not detected at the time of enlistment, seem to be aggravated by the physical and psychological aspects of the training program. It should also be noted a percentage of the other agency discharges were initially referred to those agencies by the Phase Two and Three portions of AFMET. The operational figures for 1976-77 are presented in Table II.

As can be seen from Table II, the AFMET program was successful in identifying approximately one third of the MHC recommended discharges for men and a little under one fourth for women. The low rate of identification for the month of June and part of July 1977 represented suspension of the HOI testing during the move of the Human Resources Laboratory computer from Lackland to Brooks Air Force Base. The total of eighteen hundred twenty-five Mental Hygiene recommended discharges from basic training, which included AFMET figures, represent approximately two percent of the total basic training population. It is to be remembered that these are only recommendations and that the squadron commanders have the final say in whether the trainee gets so recommended. However, once a trainee is recommended for discharge, it is a rare occurrence when the commander does not concur with the recommendation. The total number recommended by Mental Hygiene represents approximately 23% of the total number removed from the training environment before completion of basic military training by all sources during the period from 1 October 1976 to 30 September 1977. Given that most trainees are identified for removal from training at the tenth day of training, the five hundred twenty-one so recommended through the AFMET program represent a savings in training cost alone of almost one million dollars. Other costs to the government for such things as inpatient hospitalization for psychiatric reasons, maintenance of personnel on a temporary duty retirement list for psychiatric reasons, and general level of expenditures for subsequent training can only be estimated at this time. It is to be noted, however, that each trainee represents a training cost to the

TABLE II
MHC AFMET STATISTICS
Oct 76 to Sep 77

MONTH	PT COUNT	# OF MHC REC	# MHC/ AFMET REC	# & % OF MALES	# & % OF FEMALES	TOTAL AFMET IDENTIFIED
OCT	1332·	168	48	43/134 32%	5/34 15%	53
NOV	4705	150	49	46/117 39%	3/33 9%	56
DEC	7202	193	55	43/141 30%	12/52 23%	82
JAN	8143	140	53	46/118 39%	7/22 32%	73
FEB	6286	164	46	42/148 28%	4/16 25%	86
MAR	8192	179	51	41/148 28%	10/31 32%	77
APR	6782	161	47	38/141 27%	9/20 45%	72
MAY	6379	123	46	35/ 87 40%	11/36 32%	72
JUN	8249	142	12	9/105 9%	3/37 8%	25
JUL	7956	152	30	28/134 21%	2/18 11%	71
AUG	8680	119	41	35/ 92 38%	6/27 22%	94
SEP	8206	134	43	35/102 34%	8/32 25%	96
TOTALS FY77	82112	1825	521	441/1467 30%	80/358 22%	857

government of ten thousand dollars once he finishes tech school. Utilization of that estimate would suggest a total training cost savings to the Air Force of nearly six million dollars in training cost savings by the elimination of high-risk airmen early in basic training.

The future format of the AFMET program will probably be influenced by current research being conducted by members of the clinical staff and Human Resources Laboratory personnel. At the present time, data are being analyzed from fourteen thousand Air Force personnel tested last year at various AFEES on a 50-item questionnaire derived from the original 100-item HOI. Thirty-four of these items are the same as the HOI utilized by the AFMET program. Depending upon the results of this study, this version of the HOI may be given at the AFEES as part of the regular screening procedures. Due to the similarities of the two instruments, it is apparent that the HOI testing in basic training would probably need to be replaced by some other test or tests which will prove effective in further identifying low potential or "at-risk" trainees. Efforts are now being directed to locating and validating such instruments.

REFERENCES

Bloom, W. *Manual, Bloom Sentence Completion Test; Student and Adult*. Effectiveness Training Associates of San Antonio, 133 Twinleaf Ln., San Antonio, Texas 78213, 1975.

Bloom, W. Air Force Medical Evaluation Test, *Medical Service Digest*, 1977, XXVIII, No.2.

Bloom, W. Relevant MMPI Norms for Young Adult Air Force Trainees, *Journal of Personality Assessment*, 1977, 41, 505-510.

Bucky, S. F., & Edwards, D. The Recruit Temperament Survey (RTS) as it discriminates between psychoses, neuroses, and personality disorders. *Journal of Clinical Psychology*, 1974, 30, 195-199.

Guinn, N., Johnson, A. L. & Kantor, J. E. Screening for adaptability to military service. Report No. AFHRL-TR-75-30, Air Force Resources Laboratory, May 1975.

LaChar, D., Sparks, J. C. & Larsen, R. M. Psychometric prediction of adaptation for USAF basic trainees. *Journal of Community Psychology*, 1974, 2(3), 268-277.

Plag, J. A. Pre-enlistment variables related to the performance and adjustment of Navy recruits. *Journal of Clinical Psychology*, 1962, 19, 168-171.

COMPARISON OF NAVY OFFICER AND CIVILIAN PERFORMANCE EVALUATIONS

W.H. Githens and Richard S. Elster
Navy Personnel Research and Development Center

Analysis of the respective administrative systems for promotion reveals significant differences between Navy officers and general schedule (GS) civilians. Differences occur in (1) the specificity and location of the job to be filled, (2) extent to which information in addition to that on the performance evaluation form enters the system, (3) the number of people to be considered, and (4) the basis for initiating administrative action. These differences place different requirements on the performance evaluation that is to be used and the form for collecting it. It is concluded that using a form similar to the officer fitness report for civilians would be inappropriate unless the entire civilian promotion system was also changed.

As part of a Navy Material Command (NAVMAT) effort to develop methods and means for improving organizational effectiveness, the authors were tasked with a project to design a personnel development system for Navy professional civilians. This project lead directly to concern over the performance evaluation of professional civilians, for how could "development" be assessed without having some good measures of performance?

Discussions were held with civilian and military managers during which traditional material was presented concerning the differing requirements involved in (1) assessment for developmental purposes and (2) assessment for administrative purposes (promotion, demotion, merit increase, etc.). Much interest was found among the NAVMAT military personnel concerning the evaluations of civilians. General dissatisfaction was expressed over the civil service "outstanding-satisfactory-unsatisfactory" rating scale. An assertion heard many times was that a form like the fitness report used on officers is needed for the civilians. This assertion led to an examination of the promotion systems (both military and civilian) within which the rating forms are imbedded. Table 1 presents the results of this examination.

TABLE 1

COMPARISONS BETWEEN NAVY MILITARY (OFFICER) AND NAVY CIVILIAN (PROFESSIONAL) PERFORMANCE EVALUATIONS

Military	Civilian
System Differences	
1. Promotion based on # of rank needed	1. Promotion based on job to be filled
2. Fitness Reports primary information source	2. "Appraisal for Promotion" system provides additional information
3. Promotion is to a variety of jobs	3. Promotion is to a specific job
4. Decisions are periodic and made in Washington, D.C.	4. Decisions are not periodic and are usually initiated by the local supervisor
5. Many people are usually considered for promotion at the same time	5. Relatively few people are considered for promotion at the same time
6. Many raters over career	6. Few raters during career
Rating Form Differences	
7. Individual rated is compared with peers	7. Individual rated is compared with job requirements
8. A standard form is used	8. A variety of forms is used
9. Contains a number of rating levels	9. Usually contains 3 levels: O, S, and U

Due to legal constraints on the ratios of the various ranks within the military, promotion becomes based on the number of vacancies of a certain rank that occur. The work to be accomplished is organized to have jobs to be filled by the various ranks. Should there be a shortage of a particular rank at the time when a replacement is needed, a position can be filled by a lower or higher ranking officer. Rather than being driven by the ratio between ranks, the civilian system is driven by the evaluation of the positions. If a position is evaluated as being at a GS grade level, it cannot be occupied by a lower graded civilian for an extensive period.

At the time of promotion decisions, the fitness report is the primary source of job performance information on officers. The primary source of performance information for civilians is obtained by an "appraisal for promotion" system in which specific qualities needed in an incumbant of a specific job to be filled are determined. Then current and past supervisors of an applicant are asked to evaluate the applicant on these specific qualities. The performance information gathered annually on civilians is used only to indicate if the individual was outstanding, satisfactory, or unsatisfactory on previously held jobs.

Civilians are selected for promotion based on information related to how well qualified they seem to be for a specific job to be filled. Officers are not selected for promotion in order to fill a specific job. There is a great variety of jobs that may be held by officers of any particular rank.

Decisions concerning promotion of officers are made periodically (usually annually). All those meeting certain time in grade requirements are considered by a promotion board meeting in Washington, D.C. With civilians, there is no "periodic" consideration for promotion. Civilians must apply to be considered for promotion and this occurs ony when a vacancy at a higher GS grade exists or is anticipated. The selection panel is typically held in the same city as the job location.

A great burden is placed on the military promotion board in terms of the large numbers of people to be considered. Civilian promotion panels have relatively few people to consider.

Many more differences can probably be identified, but we will mention only one more. Navy officers have the same reporting senior for a shorter time than do most individuals in Civil Service. Therefore, a Navy officer in general would be rated by more raters than a Civil Service employee. When averaged, the impact of this is to level out differences in ratings due to personal rating styles of different raters. The presumption should be that the aggregated ratings of a given Navy officer would yield a truer view of that officer than would the aggregated ratings given a Civil Service employee.

The above mentioned differences in promotion systems call for differences in the forms used for evaluation and these differences in forms do exist. A very basic difference is in the "rating standard" that is used. The Civil Service forms and procedures compare the individual with the requirements of the job directly, and comparisons among individuals can be made, albeit indirectly. Navy officers are compared directly with their peers. It is assumed that they can satisfactorily fulfill the job requirements of whatever job they may be assigned at a high rank. The "compared to job standards" type of rating for civilians is not an administrative prerogative; it is required by law. In fact, it is currently illegal to evaluate civil servants on a "compared to others" basis. (Chapter 43 of Title 5, U.S. Code.)

Because many officers' records typically have to be reviewed, the Navy requires a system of performance evaluation which expedites the record, review, and decision process. Hence, a standard form like the Fitness Report seems to be mandatory for the Navy. A standard form for all Civil Service employees is not mandatory because most administrative decisions about Civil Service employees are made at the local level. Therefore, different locations may use different forms to obtain evaluative information in selecting for promotion. The rating content or topic is the same throughout the entire uniformed Navy. Variations do exist in the civilian system for the qualities to be evaluated are determined locally and the respective forms reflect this.

The civil service, "O, S, U" form is used to identify three levels of performance. This provides all the data needed for most administrative decisions, but is recognized as not providing enough information for promotion decisions. The "appraisal for promotion" system is therefore usd to obtain the needed information. The form for officers allows up to seven levels of performance (compared to peers) to be differentiated.

A benefit accruing from the situation in the Navy whereby ratees and raters are moved through various jobs is that the raters of Navy officers develop a broadly based reservoir of rating experience. Raters of Navy officers should become more astute performance evaluators (if comparison base and practice matters) than do most raters of Civil Service employees.

Major differences between civilian and military evaluation procedures exist in the average number of different raters who have rated the individual, the remoteness and massiveness of the administrative decision-making apparatus, and the standard of comparison to be used when making ratings.

The civilian rating form and the officer Fitness Report seem geared to the system within which they are used. Comparison of these systems has revealed many basic and significant differences. It is concluded that using a form similar to the officer Fitness Report on civilians would be inappropriate unless the civilian promotion and administrating system was also changed accordingly.

REFERENCES

United States Code Chapter 43, Title 5

Federal Personnel Manual Supplement 990-1, Instruction 136, Part 430, August 23, 1973

Federal Personnel Manual Chapter 430, Instruction 198, August 23, 1973

Federal Personnel Manual Instruction 242, May 25, 1977

Office of Civilian Manpower Management Instruction 430C, January 28, 1972

Department of the Navy, *Bureau of Naval Personnel* Instruction 1611.12, January 10, 1978

SYSTEM PSYCHOLOGY APPLIED TO PERFORMANCE ASSESSMENT OF USAF CIVILIAN EMPLOYEES

By

Dell C. Toedt and Forrest R. Ratliff
Manpower Development and Evaluation Branch
Personnel Research Division
Air Force Human Resources Laboratory
Brooks Air Force Base, Texas

In January 1977, the Air Force Human Resources Laboratory (AFHRL) received a Request for Personnel Research (RPR) from the Directorate of Civilian Personnel, Headquarters USAF (AF/DPCM), requesting a research effort which would provide a new system for evaluating civilian employees. The RPR was evaluated by the Personnel Research Division, Personnel Manpower and Development Branch, which concluded it was feasible. In this study, it stated that it would take 3 months to develop a research plan, and this was briefed to Air Force Headquarters in May 1977. It was the position of AFHRL at that time that some additional professional staff would need to be added to develop the research plan. The proposal was reviewed, and AFHRL received approval to begin work early in September 1977.

RPR 76-40, *Supervisory Appraisal of the Current Performance of Air Force Civilian Employees*, is a very well written statement of a customer requirement. It was written by Mr. Charles J. Hugsa, now retired from civil service, and contained clearly defined problem areas and objectives for AFHRL to work for. Keep in mind that Ratings refer to specific job performance while Appraisal includes promotion potential. It stated that the Air Force civilian appraisal program is losing its effectiveness to differentiate among employees with respect to their current performance because element scores have become inflated after several years. The deterioration of the system was characterized by one or more of the following features.

1. There was a clustering of element scores in the two top points on the scale.

2. There had been a sharp increase in the average element score together with a sharp reduction in its standard deviation.

3. Larger numbers of employees received maximum scores on all elements on which they were rated.

4. Average element scores tended to increase as the skill level or grade increased.

5. The degraded appraisal instruments were being used for a variety of purposes--selection for training, details, reassignment, merit promotion--and this led to loss of effectiveness for achieving any of those puproses. (Remember this point when you read the comments on the Uniform Guidelines for Employment Selection Procedures.) There was an ultimate loss of credibility to management, employees, and supervisors.

6. Several different appraisal systems are being operated simultaneously to meet different requirements, e.g., the Performance Rating Act of 1950, the Incentive Awards Act of 1954, and the Salary Reform Act of 1962, as well as others.

A section of the RPR was devoted to state-of-the-art as seen from Headquarters USAF. This was the most understated section of the RPR as it briefly covered some Civil Service Commission directives, Air Force policies, and practical considerations from an economy standpoint. It did not mention the Civil Rights Act of 1964 or any of the complications influenced by Equal Opportunity, Affirmative Action, Public Information Act, Privacy Act, or the far reaching court decisions by which they were implemented.

Three objectives were defined for this effort.

1. Provide management with multi-purpose instruments to obtain information on employee performance necessary to make sound decisions affecting the work and careers of Air Force civilian employoees.

2. Provide rating supervisors with sound information on which to make the determinations required by the Performmance Rating Act of 1950 (5 USC 43).

3. Restore the credibility of the Air Force appraisal program in the eyes of all concerned.

One of the less recognized by-products of the new national awareness of emphasis on human performance, Civil Rights and EEO compliance, design to cost, or any of the many people oriented programs is that the demand for psychologists has resulted in a shortage of psychologists with experience in large, people dominant systems. Early on, it was found that two authorized Air Force "blue suit" psychologists were not available, so temporary civilian positions were substituted. There is a talent search continuing for research psychologists who are experienced in this area.

The most controversial Air Force personnel program which has had a research beginning in past years has been the new Officer Efficiency Rating (OER). That represented an effort to rate the performance of approximately 125,000 college trained, professional level management, and technical and scientific "white collar" people. This new effort for civilians has shaped up to be bigger, broader, and far more complex than the OER program. There are several reasons: Civilian employees are categorized by classes of occupation; Classified workers are grouped and identified in categories of jobs sufficiently homogeneous as to subject matter, level of difficulty and responsibility, and qualification requirements to warrant similar treatment in personnel policies and pay administration; General Schedule (GS) workers are divided into 18 grades and paid on a nation-wide scale adjusted annually by the President under rules established by Congress; Direct Hire workers who are in the skills and crafts are known as Wage Grade employees and are paid on a scale that varies by prevailing wages in the local area. That diverse occupational spread makes tyring to rate civilians on one rating scale similar to rating all officers, non-commissioned officers, and airmen on one system. The scope of the civilian force is shown in Figure 1.

FIGURE 1

Number Of Civilian Employees By Pay Plan

ANNUAL PAYROLL	WS, WG, WL		GS	ANNUAL PAYROLL
3,000,000	303	1	124	800,000
11,000,000	1,264	2	1,465	11,000,000
12,000,000	1,159	3	9,935	85,000,000
29,000,000	2,925	4	16,292	94,000,000
51,000,000	4,978	5	19,685	213,000,000
66,000,000	5,923	6	7,075	85,000,000
82,000,000	7,035	7	10,654	139,000,000
103,000,000	8,692	8	2,647	39,000,000
134,000,000	10,284	9	15,457	247,000,000
335,000,000	25,269	10	990	17,000,000
95,000,000	6,669	11	14,705	287,000,000
43,000,000	2,793	12	12,584	293,000,000
13,000,000	773	13	7,540	221,000,000
7,000,000	373	14	2,729	90,000,000
3,000,000	126	15	877	34,000,000
1,000,000	47	16	79	4,000,000
350,000	14	17	9	500,000
50,000	2	18	5	240,000

USAF Advanced Personnel Data System - Civilian, Oct 77

In 1887, when the rating system was installed in the War Department, almost any factor or technique could be used, but how a person performed his current job was considered to be the best predictor of future performance and that was the basis for promotion. The system has undergone periodic change since that time, and every rating system has lasted 5 to 10 years before it was rendered inoperative by participant "gaming" or by changing social conditions. Changes developed by management, administrators, and by "psychologists" broadened the rating effort across traits, personality and even as far as psychoanalytic theory. Recent legislation and court decisions have brought us back to full cycle. Based on recent legislation and court interpretation, all selection, training, and promotion again will be based on job related factors—"EXCEPT"—and it is the except that makes it interesting.

Although there has been little criticism of the intent of the Equal Opportunity and Civil Rights legislation, the same cannot be said for the proliferation of directives coming from interested federal agencies. Since November 1976, two different, and apparently conflicting, sets of guidelines have been issued. The *Guidelines on Employee Selection Procedures* of the Equal Employment Opportunity Commission (29 CFR Part 1607) contains one set of directives issued as a minority report because they would not accept the *Federal Executive Agency Guidelines* issued jointly by the Civil Service Commission, Dept of Labor, and Justice Department. On 30 December 1977, a proposed rule-making titled *Uniform Guidelines on Employee Selection Procedures* was published in the Federal Register which has the stated intent to jointly issue one guide for all. The result is basically the EEO Guidelines although they have not excluded the worst parts of each. The lack of internal consistency and specificity has resulted in a position that should be of grave concern to everyone engaged in any aspect of rating, appraisal, or any aspect of ANY employment decision such as hiring, promotion, demotion, membership, referral, retention licensing and certification, selection for training, transfer, or any other decisions related to the above.

Some isolated examples are use of terms such as "relevant labor market," "work behavior", and reasonable effort" being used to define specific requirements. Internal consistency is lacking between some of the policy statements and validation requirements, or Affirmative Action Programs.

Highlights of the current Civil Service directives preclude withholding information from employees on appraisals, ratings, or any other information which concerns their job performance. It also prohibits using a "single" total score for each employee and that scores must be derived from those elements in an appraisal form that are relevant to the position to be filled. Another prohibition is that no "personality" type variables can be used to describe an individual. With these (and many more) contratints, the Lab cautiously approached what turned out to be a new phase of research for us.

Traditional psychological research has been laboratory oriented, with a technical report published at the completion, and the results of the data analysis furnished to the customer for his use as information to be used in solving his problem. At the beginning, there was a distinct but sincere professional difference of opinion in the recommended approaches to this problem by the professional staff. Those psychologists who had been accustomed to traditional academic research recommended that these classic principles were the best approach

to this problem, while the staff who were systems oriented championed a systems oriented, customer related approach based on research but designed to produce a prototype rather than a report. This was conceptually difficult because there was no precedent for applying research on the systems approach in the civilian Department of the Air Force sector. It was finally resolved in favor of a systems approach, however, doing research in the field using Air Force civilians and in partnership with the Directorate of Civilian Personnel. The author of the original RPR retired, so Civilian Personnel secured a replacement psychologist to represent their interests and to help in the translation of research requirements to personnel system characteristics and to coordinate the field work. The effort was separated into two sections: one was the systems development side, and the second was the research program. This project report outlines the systems development up to this time, and the research efforts will be repoted on by Lt Colonel Forrest Ratliff in a presentation to follow this one.

The derived approach to solving the customer's problem was a dual approach combining laboratory and field research and, from that, developing a prototype system based on our findings; then to conducting an operational test and evaluation (OT&E) in the acutal environment and testing for validity and relevance. During the OT&E, we plan to develop the necessary linking to the existing personnel management systems so that our prototype will both interface and support existing management oriented systems. This, in itself, is not a small problem because millions have been invested in data processing capabilities which are too expensive to replace to accommodate some new revolutionary system. This posed several problems for the psychological side because it meant becoming familiar with a large, complex management system, and it emphasized the concept of concurrency and relevance as an "add-on" rather than a completely new development. This concept of customer oriented research limited our approach to a practical, usable prototype system as opposed to a theoretical new concept that might be developed out of context of the real world.

There are several novel and exciting aspects to this effort. Fundamentally, it is because the civilian side of the military is the least understood component of the defense effort. The Committee on Armed Services of the 95th Congress noted that few institutions have been as widely criticized as the Federal Civil Service. The fact that civilians constitute over 20% of the Defense Department, and have since World War I, has made it difficult to analyze their impact on the defense effort. Traditionally, military analysts have shown far more interest in developing new weapons systems or tactical and strategic systems than in the support structure in which civilians play such a vital role. Projected manpower shortfalls for recruiting and the skyrocketing costs of manpower bring the contributions of the civilian sector into sharper focus.

We found that doing research on civilians is relatively new to the Air Force. There are no established procedures, and since civilians come under Title VII of the Civil Rights Act the established military procedures are not completely appropriate. There are conflicting guidelines about what can and can't be done and, in addition, unions represent a new organizational entity which must be considered. A very recent change to the Federal Personnel Manual directs that unions participate in developing new appraisal systems. The conflicting provisions of the Feedom of Information Act and the Privacy Act have not been fully tested in the courts, particularly as pertains to research and the needs for data in attitude and aptitude measures. The increasing awareness of the civilian work force to individual rights and the apparent willingness of some employees to exploit them pending future clarification cause time consuming deliberations about how to proceed, particularly in the face of lack of good evaluation data. These efforts are breaking new ground in the area of personnel research within the Department of Defense, but both the procedures and the findings should be of interest to the psychological community as a whole.

One last aspect to systems development is our proposed effort to establish a modest follow-on testing procedure in order to monitor the effectiveness of the system and to identify system deviations from "gaming," which we know to be inevitable, or from basic changes in the organizational characteristics which may influence the appraisal effectiveness of our proposed system.

Finally, as the prototype system is refined in the field and with the cooperation and participation of the customer, we will begin the turnover phase in which it will be implemented throughout the Air Force. The personnel in the field who have been trained in administering the research and testing efforts will form the nucleus of the implementation effort, and we intend to turn the prototype over with minimum disruption to the customer's work effort.

The documentation of our research will be oriented to future use of the psychological community, but the findings will be reported to the user in customer oriented format. It will include significant reports. These will be consolidated into the final implementation recommendations for USAF DPCM.

This has been an interim project report. It is the beginning of an effort which has the potential to shape a part of the Air Force for the next 10 to 20 years. It is an ambitious project; one that will directly affect more people than any large scale psychologically based system that we are aware of. It is the only large personnel management system to be designed, developed, and tested within the psychological community, and we are proud to be part of it.

BIBLIOGRAPHY

Campbell, J.P., Dunnette, M.D., Lawler, E.E. III, Weick, K.E. Jr., (eds.) *Managerial Behavior, Performance, and Effectiveness.* New York: McGraw-Hill, 1970.

Defee, D.T., Management by Objectives: When and how does it work? *Personnel Journal,* 1977, *56,* 37-42.

DeGreene, K.B. (Ed.). *Systems Psychology.* New York: McGraw-Hill, 1970.

Federal Register. *Uniform Guidelines on Employee Selection Procedures.* Vol. 42, No. 251.

Holley, W.H. Jr. Performance appraisal in public sector arbitration. *Public Personnel Management,* 1978, *8,* 1-4.

Johnson, C.A., Meehan, J., & Wilkinson, R.E. *Officer effecitveness report development - 1971 through 1972.* AFHRL-TR-76-61) Lackland Air Force Base, Tex: Personnel Research Division, Air Force Human Resources Laboratory, September, 1976.

Koontz, H., *Appraising managers as managers.* New York: McGraw-Hill, 1971.

Management By Objectives Key Operating Indicators. *Performance Appraisal and Planning Guide.* Federal Home Loan Bank Board. Washington: 1977.

McDonald, J.C., *An historical survey of performance rating systems in the federal service*. U.S. Civil Service Commission, Washington: U.S. Printer's office, 1974. (NTIS No. PB-258 239)

Odiorne, G.S., *Management by Objectives*. Belmont, Calif: Pitman, 1965.

Pearman, J.R. A systems approach to personnel management: Advantages and disadvantages of implementation. *Public Personnel Management*, 1978, *8*, 6-18.

Schoderbek, P.P., & Plambeck, D.L. *The missing link in management by objectives-continuing responsibilities*. Public Personnel Management. 1978, *8*, 19-26.

Shaping the defense civilian workforce. Report No. 96-743, Washington: U.S. Government Printing Office, 1977.

Stanton, R.R. Professional managers and personal liability. *Public Personnel Management*, 1978, *8*, 43-48.

United States Civil Service Commission. *Fairness of psychological test: Implications of four definitions for selection utility and minority hiring*. Washington: U.S. Government Printing Office, 1975 (NTIS No. 8B-258 215)

United States Civil Service Commission. *Fairness of selection tests a critical analysis*. Washington: U.S. Government Printing Office, 1976. (NTIS No. PB-261-699).

United States Congress, Senate. Committee on Armed Services. 95th Congress, 1st Session, Washington: U.S. Government Printing Office, 1977.

Varney, G.H., *Management by Objectives*. Chicago: Dartnell, 1971.

RESEARCH IN SUPPORT OF A CIVILIAN APPRAISAL SYSTEM

By

Forrest R. Ratliff, USAF, and Dell C. Toedt
Manpower Development and Evaluation Branch
Personnel Research Division
Air Force Human Resources Laboratory
Brooks Air Force Base, Texas

My purpose here is to describe the research program now underway in AFHRL/PE to meet the objectives of RPR 76-40, the development of a DAF civilian personnel appraisal system. In addition, I would also like to share with you the rationale, logic, and assumptions that provided the foundations for the current research program. These latter may be more important in one sense than the specific elements of the requirements document inasmuch as they tend to influence the content in many ways, depending upon motivation, objectives, and other intangibles.

Mr. Toedt referred to an array of specific requirements, to the judgmental process leading to a systems approach, and to the internal dialogue concerning whether research objectives should focus only on scientific reports or upon the broader problem of delivering an operational system. The actual situation was much more complex, with a number of processes influencing and contributing to our perception of possible solutions and various models of research support.

Initially, we had a large, basic research program focused on some of the fundamental problems in the performance evaluation area. This research included research on rating variance, synthetic validity, proficiency testing, and the dimensions and effects of the rater-ratee interface. My co-worker, Dr. Cecil Mullins, had developed hypotheses and evidence over the years: (1) that some raters are more accurate than others; (2) that ipsative ratings (i.e., ratings of traits within an individual) may be useful in personnel operations if reliable anchor variables can be found; (3) that synthetic validity, a structured decision model for validating selection procedures by predicting training outcomes from aptitude test scores, might be further developed; and (4) that job knowledge tests which exploit experience as well as training might be useful in criterion work. I had been working on the concept that well defined rating factors may provide a major key to improved ratings and assessments. I sensed that many rating systems may overburden the rater with discriminations that are not properly in the rating or perceptual judgment domain. Rather, a certain amount of inaccuracy or unreliability might be forced upon the rater by the inappropriateness of casually assigned or unevaluated judgment facotrs.

One of the major problems in this area is the lack of objective or reliable job performance evaluation methods or, if you will, the long sought after job criterion. The validation of the findings from research on ratings or other methodologies ultimately depends upon the availability of some form of objectivity in the measurement of job behaviors. In reviewing the literature, I looked for those studies concerned with the interface between the supervisor and the worker. I felt that if a criterion existed for jobs which lack concrete end products or repetitive production tasks, then it must be made up of the degree of agreement between the perceptions of the supervisor and the worker about the job, i.e., what it is, how it should be done, the responsibilities and requirements which must be met, and other factors. The extent to which they (the rater and ratee) agree upon the dominant elements of the job or share common perceptions of the job determines to a large extent whether a useful or reliable criterion can be established for measurement purposes. There is little doubt that it exists in some form. Unfortunately, very few research studies are available on this topic in the literature. The apparent lack of significant entries in the research literature on supervisor worker job perception—a highly strategic arena for defining and obtaining criterion information—increased the drive to structure this area for research.

You may note from the above description that a strong interest in the criterion problem existed, the experience in developing a performance evaluation research program was there, and the willingness to accept a challenge was there. Thus, when the RPR arrived, we were prepared to respond positively.

However, we did not recognize at that time the complexity of designing an appraisal program to meet the guidelines in today's socio-legal climate. The problem was initially considered resoluble through research on a potential-opportunity-energy construct proffered by Dr. Mullins. However, problems in the reqiured testing program were perceived. At this point a necessity to develop a research program that was much broader than the intended appraisal instrumentation was recognized. This broad program was needed to provide a range of options for coping with the many known hazards in providing assessment methodology. Also, USAF DPC began to recognize the need for some form of final product which could be implemented in the appraisal program. This research concept was briefed to USAF/DPC and approved with direction to provide a detailed research plan. It was during the development of this plan that the systems approach described by Dell Toedt was formulated for synthesizing both research and development elements into a coherent appraisal program.

The research program to support the development of a new civilian appraisal program was developed in two phases. The first phase was developed in a briefing to USAF/DPC in which a research approach was the primary concern. The research approach was developed from considerations based on the complexity of the problem, the need for objective measures of job performance, and the nature of the appraisal process which was regarded as deeply meaningful to the individual, and which should be institutionally related, objective, and with multiple sources of input. This basic approach, with prototype examples of planned research, was approved by USAF/DPC with instructions to provide a detailed reserach plan within 3 months. This report describes the research plan which was presented to USAF/DPC as the basis for the development of a new civilian appraisal system. As the problem of developing an appraisal methodology was studied in the context of management needs, individual rights legislation, and court decisions, it became obvious there were several driving forces which had to be reckoned with. These forces generated requirements and criteria which affected the basic design and scope of the entire project. These requirements were job relevance, variance based on merit, credibility and acceptability with both management and workers, and compatibility with evaluation concepts across all the systems elements. In addition, th Request for Personnel Research (RPR) 76-40 required that at least two measures be developed, one reflecting job performance, and the other assessing promotion potential.

The research program was structured to provide the means to remedy gaps in the assessment state-of-the-art, to modify methodology as needed, and to develop required new approaches. The research program and the systems development approach described earlier are viewed as synergistic in providing the methodology and data for operational determination of job performance, promotion potential, training needs, assignments, and special details. We believe that the development of a system which does deliver the above evaluation products with reasonable objectivity and reliability will develop credibility/acceptability with both the work force and management.

The completed research plan described three major components. These were (1) a job performance measurement module, (2) an appraisal or estimate of potential module, and (3) an operational test and evaluation (OT&E) module. In addition, there were a number of research units designed to feed information and methodology in these modules. A large component of rating research is underway to provide needed advance in the state-of-the-art. Two separate measures of job performance and estimates of potential were required by the Civil Service Commission. The OT&E component was designed to provide (1) qualitiative and quantitative information about components of the system, (2) an integrative function in the assembly of the final components of the system, (3) the development of an operational data base for evaluating the dynamics of the system in a realistic environment, and (4) the general analytic context for developing an appraisal system monitoring and updating package.

The job performance measurement module is oriented around the concept of supervisor/worker perception of the major elements of the job. The behavior model requires the worker and supervisor to agree upon the kind and magnitude of tasks to be accomplished over the next quarter to be used as the basis for evaluating performance. This approach is derivative from the rater-ratee interface construct and the management by objectives concept. We call this approach an Evaluation by Objectives Technique. The formal work units contained in this module are (1) evaluation by objective (EBO), (2) rater-ratee interface (R-R), and (3) acceptability/credibility (A/C) of the appraisal program. The rater-ratee interface research should have direct inputs into the EBO module as development progresses. The acceptability/credibility research is designed both to demonstrate compliance with USAF/DPC'S requirement that the system is a fact more acceptable/credible to the work force and to identify these components/functions/attributes of the appraisal process which are meaningful and accepted by the work force. To date, there is a very little research in the literature on this problem, and it is, I believe, a substantive one.

The second module, providing an estimate of potential for promotion or other assignment, is basically conceptualized as a prediction problem. That is, the task is to predict from among those available (1) who would be most successful in a new job assignment, and (2) the performance of the individual in a specified capacity. Prediction of this type represents a somewhat different problem than that involved in the measurement of performance. For the prediction problem, a rather extensive data base on the interrelationships of aptitudes, ratings, background or experience factors, and job performance is required; and, along with some algorithm for using these elements to predict a reasonable criterion.

We are looking to the job performance measurement module to supply the criterion reference marks for the predictive aspects of the appraisal program. To integrate the rather large masses of data into useful weighted, standardized indexes, we hope to apply techniques from policy specification techniques as developed by the laboratory.

I would like to express some of our views on ratings and intentions for their use in the personnel system. Ratings are viewed as an integral element of almost every personnel evaluation; system. A great deal of research has been performed on ratings, and the bibliography in the current literature runs to hundreds of titles. The basic problem driving the needs for further research on the use of ratings is that some very cognizant issues have remained practically unaddressed and some assumptions have been accepted with inadequate evidence about a rater's ability to rate. The questions we are addressing in this area deal with the influence of rating factors upon rating behavior and the questions pertaining to rater accuracy. The difficulties of obtaining merit related variance will be examined in the context of the new appraisal program by use of a spread variance technique, and by the use of feedback to the supervisor on his ratings, and on the desired evaluation goal. Holding the supervisor accountable for his ratings by using them in assessing his supervisory ability is under study and will be implemented in OT&E. In addition, the potential for producing variance which is inherent in ipsative ratings (that is, the ranking of traits or factors within the individual) is being given a new look--a very, very close scrutiny in fact--to determine if the ipsative rating concept can be applied in a personnel system.

The problem of using ipsative scores has been the lack of a means for making a comparison across individuals. The possibility of developing an anchoring variable which would be useful for developing a normative base for ipsative ratings is under active investigation. In addition, there are studies on the utility of ipsative ratings on both individuals and jobs.

The third research module is concerned with the evaluation of our research products in the field. We are committed to the delivery of an operational prototype job performance measurement and personnel appraisal system. This commitment is quite broad and assumes that an operable prototype appraisal system can be developed and assembled for tryout on a representative sample of Air Force bases. The key element in this goal, after the development process has been completed, is that a significant effort must be made to evaluate the operating characteristics of the components of this proposed system in a realistic environment. This involves developing the methodology, constructing operational data collection channels, designing the data base and analytic constructs for estimating the reliabilities of the evaluation system components, and for determining the psychometric properties of the appraisal elements for predicting job success as is now required by the courts. The system assembled for implementation from the integrative research units will be subjected to a regorous operational test and evaluation (OT&E) in the most realistic environment obtainable under current USAF DPC policies and guidelines. There are few prototypes for this evaluative approach in the personnel domain, and we are, in fact, moving into relatively uncharted territory with this enterprise. The successful development and application of the OT&E concept to a population as large and complex as the DAF will require an immense amount of prior planning, pretesting, and procedural development if objective data generated in a realistic appraisal environment are to be obtained. The problems associated with realistic environmental testing in the personnel systems domain have scarcely been surfaced--let alone defined, analyzed, and fitted into a set of operational assumptions which will carry the load of both rigorous statistical evaluation and the need to adapt to the many constraints inherent in the operating personnel systems of today.

This new system must be responsive to a number of regulatory, legislative, and court imposed constraints, and yet achieve management goals for a fair and equitable appraisal which each employee rightfully feels is his due. The extent to which we are successful in this undertaking will determine the value of our enterprise to both the worker and management.

We have described a complex, comprehensive program for developing a civilian appraisal system in a large, dynamic organization. We realize that in taking the described approach we are breaking with many precedents and are aware of some of the dangers involved.

This system has been conceived as a reasonable approach to the problem of designing a new appraisal system from scratch, and it is not viewed as the ultimate in such systems. Rather, it appears to be the minimum to do the job. We will be very satisfied if the system works and produces acceptable reliabilities and validities and are optimistic that those modest objectives can be reached.

Some of the individual efforts described in the plan are acknowledged as high risk. However, we do believe that the approach developed is conceptually strong and represents a practical, theoretically sound approach to the general problem of performance evaluation in large organizations. We are confident this systems apprach is addressing the fundamental issues and will influence the basic pattern of future thinking and research about appraisal systems.

EVALUATION OF THE QUICK TEST AS A SCREENING DEVICE

Stanley H. Holgate and Linda E. Bailey
Biomedical Laboratory
Aberdeen Proving Ground, Maryland 21010

The Quick Test (QT) was designed to provide quick screening of verbal-perceptual intelligence in applied settings (Ammons and Ammons, 1962). Of the two reviews in *The Sixth Mental Measurements Yearbook*, one suggested that the QT was worthy of further research as a clinical tool (McCandless, 1965), and the other relegated the QT to the role of a large-scale screening test when time is of the essence (Piers, 1965). Still, the QT IQ scores correlate well with the WAIS Full Scale IQ (Joesting and Joesting, 1972), but generally achieves the highest correlations with the Verbal Scale IQ and the Vocabulary subtest of the WAIS (Feldman, 1968; Cull and Colvin, 1970). When compared with other tests, the QT generally correlates better with verbally related (vocabulary, reading, spelling) tests than with arithmetic of performance measures (Whitney and Metzger, 1965; Mednick, 1969).

In the process of selecting human subjects for many studies it is desirable, if not necessary, to match the subjects on the basis of measured intelligence. Normally, sufficient time is not available to administer the WAIS or even one of its short forms. Furthermore, the available manpower may not be adequate to administer the QT according to the standard instructions.

The present study was conducted to evaluate another method of administering the QT and to compare it with the standard procedure. In the process of the evaluation, the QT was compared with other measures as well.

METHOD

A total of 193 US Army enlisted men volunteered to serve as subjects.They ranged in grade from E-1 to E-6 and had a mean age of 21.63 ± 3.59 years. They were tested while being processed as participants in the Medical Volunteer Program. The data were collected over several volunteer cycles since there were only 40 to 50 volunteers in each two-month cycle.

Three groups of subjects (Groups A, B, and C) were given either Form 1, 2, or 3 of the QT according to Method I (M-I), the standard method described in the provisional test manual. The subject pointed to or named the line drawing he thought best fitted the test word given orally by the tester. The test continued until the subject made six consecutive errors on the more difficult words and six consecutive correct choices on the easier words. The raw score was the number of correct responses including the easier words not given. These same subjects were then given the other two forms of the QT by Method II (M-II). Under the conditions of M-II, the subject was given a test plate of line drawings and a corresponding answer sheet containing a set of instructions and the list of test words. The subject was instructed to write the number of the drawing which best fit each word in the blank next to the word. If the subject had no idea what the correct response was, he was instructed to write a zero in the blank. This method of testing was self-paced. The subject's raw score was the number of correct responses. The data from the initial three groups of subjects was so promising that four more groups of subjects (Groups I, II, III, and IV) were administered all three forms according to M-II.

During the processing upon arrival at the laboratory, the subjects were given the Proverbs Test (PVT) (Gorham, 1956), the Number Facility (NF) test (Moran and Mefferd, 1959), and the QT. Some of the subjects also took the Revised Minnesota Paper Form Board Test (PFB). The subjects' General Technical (GT) aptitude scores from the Army Classification Battery and their education level (EL) were obtained from the subjects' records and combined with the results of all the above named tests to form the data base described in the results section.

RESULTS

The subjects in this report most closely approximated the Ammons and Ammons (1962) adult standardization group. For the 193 subjects in this study, the QT mean combination raw score was 122.17 ± 10.50 and the median score was 123.14 which translates to a group mean IQ of about 99 according to Ammons and Ammons (1962).

The means and standard deviations for all groups on all test measures plus education level (EL) are presented in Table 1. One-way analyses of variance were computed to determine if there were any statistically significant differences among the groups. The analyses of variance revealed significant differences among groups on GT scores and EL. The differences on PFB approached significance, but as shown by the standard deviations, the large within group variance prevented finding a significant difference. Further analyses using paired *t* tests revealed that Group I had significantly higher scores on all measures except QT total (QT_T) which was the sum of the three separate forms. Further analyses of variance among the three different forms of the QT given by M-I (shown as QT_{Ei} in Table 1) or M-II (shown as QT_{Si} in Table 1) revealed no significant differences between M-I and M-II.

Correlational analyses were made separately on the data for Groups A, B, and C and Groups I-IV. Significant interform correlation coefficients were obtained for the QT irrespective of method of administration. QT scores correlated significantly with GT scores and highly, though not significantly, with PVT scores. Not surprising were the weak relationships observed between the QT scores and the NF and PFB scores. The order and size of the correlation coefficients support the view that the QT measures individual verbal ability more precisely than it measures spatial concept formation, arithmetic computation, and eye-hand coordination.

CONCLUSIONS

The findings of greatest importance was that of no significant difference between Method I and Method II in administering the QT. Thus, Method II stands on its own merits as an alternative method of administration when the QT is used as a screening device for matching experimental subjects. Method II should nonetheless be limited to the research setting and not supplant the clinical function of Method I. The QT does have limitations as a selection mechanism. It would not be of much value in matching subjects for speed tests involving spatial concepts and arithmetic operations. Furthermore, the QT is not quite as powerful as some of the other measures used. That is,

TABLE 1. Group mean raw scores and standard deviations for the experimental measures used in this study. Education level (EL) is included for groups for which it was available. Measures QT *Ei* and *Si* were given by Methods I and II, respectively.

	GROUP						
MEASURE	A	B	C	I	II	III	IV
n=	25	25	22	35	19	23	44
QT_{E1}	41.56						
	3.91						
QT_{E2}		39.12					
		4.02					
QT_{E3}			38.68				
			4.35				
QT_{S1}		40.44	39.77	42.34	41.89	41.30	40.57
		3.74	5.02	6.51	3.83	4.56	4.40
QT_{S2}	40.40		39.68	40.83	40.47	40.26	41.16
	3.48		3.58	2.93	3.24	3.53	3.82
QT_{S3}	40.84			41.43	41.42	40.48	41.09
	4.18			3.47	3.36	3.94	4.14
QT_{T}				124.60	123.79	122.09	122.86
				9.39	8.69	10.34	11.08
PFB				45.43	41.00	------	40.00
				7.21	12.45		11.02
PVT	28.32	28.40	29.82	29.46	28.21	29.96	26.98
	3.99	4.83	5.14	5.27	5.35	5.67	5.75
GT	112.44	111.04	112.91	122.40	108.84	113.84	109.34
	13.45	13.96	16.82	12.26	13.99	13.56	9.39
NF	41.56	40.12	39.86	42.60	35.95	36.74	35.84
	3.91	4.02	10.36	11.60	8.46	9.46	11.52
EL				13.06	12.58	12.78	12.05
				1.49	1.81	1.04	0.71

the QT failed to make as fine a discrimination among groups as some of the other measures used in this study. Even so, the ease of administration and scoring as well as the fact that the three forms of the QT permit followup or repeated testing do support the use of the QT in matching subjects on the basis of verbal intelligence and general aptitude.

REFERENCES

Ammons, R. B., and Ammons, C.H. The Quick Test (QT): provisional manual. *Psyhcological Reports*, 1962, *11*, 111-161.

Cull, J.G. and Colvin, C.R. Correlation between the Quick Test (QT) and the WAIS Verbal Scale in the rehabilitation setting. *Psychological Reports*, 1970, *27*, 105-106.

Feldman, S.E. Utility of some rapid estimations of intelligence in a college population. *Psychological Reports*, 1968, *22*, 23-26.

Gorham, D.R. *Clinical manual for the Proverbs Test*. Missoula, Montana: Psychological Test Specialists, 1956.

Joesting, J. and Joesting, R. Quick Test validation: Scores of adults in a welfare setting. *Psychological Reports*, , 1971, *30*, 537-538.

Mednick, M.T. The validity of the Ammons' Quick Test of intelligence. *Psychological Reports*, 1969, *24*, 388-390.

McCandless, B.R. Review of the Quick Test. In O.K. Buros (ed.) *The Sixth Mental Measurements Yearbook*. Highland Park, New Jersey: Gryphon Press, 1965. Pp 825-826.

Moran, L.J. and Mefferd, R.B. Repetitive psychometric measures. *Psychological Reports*, 1959, *5*, 269-276.

Piers, E.V. Review of the Quick Test. In O.K. Buros (ed.) *The Sixth Mental Measurements Yearbook*. Highland Park, New Jersey: Gryphon Press, 1965, Pp 826-827.

Whitney, V. and Metzger, R. The Quick Test as an intelligence screening device in a large scale program of employee applicant testing. *Journal of Clinical Psychology*, 1965, *21*(1), 71-72.

A PROPOSED SELECTION SYSTEM IN A NAVAL AVIATION

by

Richard E. Doll, MSC, USN[1]
Naval Aerospace Medical Research Laboratory
Pensacola, Florida 32508

INTRODUCTION

The content of this paper is the product of several interactions among personnel at the Naval Aerospace Medical Research Laboratory, the Chief of Naval Air Training, the Navy's Task Analysis Evaluation Group, and the Naval Training Equipment Center. The system, as proposed, is still awaiting final approval and funding.

The energy crisis, high cost of aircraft procurement, maintenance and operating costs, personnel costs, coupled with lack of training airspace, have dictated an examination of additional methods of identifying potential attrites and the selection of candidates for the various pipelines. The various written flight aptitude tests administered over the years have served as initial screening procedures, and have effectively served this purpose. However, these tests do not identify certain important psychomotor and cognitive abilities that are relevant to aviation performance.

Figure 1 presents the psychological psychophysiological characteristics that are considered important in Naval Aviation, the Navy's present measures of these traits, and the proposed measures of these traits. Presently, it is felt that the Naval and Marine Corps Aviation Selection Test Battery, consisting of the Academic Qualification Test (AQT), the Mechanical Comprehension Test (MCT), the Spatial Apperception Test (SAT), and the Biographical Inventory (BI), effectively assess the dimensions of general intelligence, mechanical comprehension, spatial apperception, and relevant biography, respectively. The trait of motivation is measured to some extent by the Biographical Inventory. This dimension has been highly resistent to all efforts and there is nothing new on the horizon, although the possibility of examining an assessment center approach is being entertained. While the present selection system measures four of the dimensions satisfactorily, it leaves five dimensions unaddressed; viz., psychomotor, divided attention, selective attention, stress/anxiety, and motion reactivity. The selection system proposed in this paper would address all but the dimension of stress anxiety. The latter is presently being investigated at NAMRL in which a voice analysis procedure is undergoing validation research. However, this effort is too preliminary and basic at this time to be considered as part of the present proposed selection system.

FIGURE 1
Psychological Psycho-Physiological Dimensions Inportant in a Naval Aviator and the Navy's Present and Proposed Measures

	Gen. Intell.	Mech.	Spatial	Rel. Bio.	P/M	Div. Att.	Sel. Att.	Stress/ Anxiety	Motion React.	Moti- vation
Present Measures:	AQT	MCT	SAT	BI	- -	- -	- -	- -	- -	(BI)
Proposed Measures:	AQT	MCT	SAT	BI	PTD	IMPACT	D/L	- - (R&D-VA)	BVDT	(BI)
OTHER:	. . . . APAMS									

The proposed selection system would address the unmeasured performance dimensions as follows:

1. Psychomotor--Use the perceptual psychomotor complex coordination test developed by the Air Force.
2. Divided Attention--Use the NAMRL developed Integrated Multi-Task Psychomotor and Cognitive Testing (IMPACT) system.
3. Selective Attention--Use a dichotic listening test similar to that being used by the Israeli Air Force.
4. Motion Reactivity--Use the NAMRL developed Brief Vestibular Disorientation Test (BVDT).

[1]Opinions or conclusions contained in this report are those of the author and do not necessarily reflect the view or the endorsement of the Navy Department.

It is intended that all of the above will be used in an 18-month pass through study to allow for a thorough analysis of validities, redundancies, cutting scores, and administrative procedures. In addition to the above trait measures, it is planned that a system, referred to as the Automated Pilot and Aptitude Measurement System (APAMS) developed at Lackland Air Force Base, be included in the study. The Navy's Task Analysis Evaluation Group (TAEG) conducted a study in which they concluded that such a system offers the potential for increased precision in selecting aviators and for reducing later attrition. It is difficult to place APAMS under any one trait category, since it approximates a job sampling approach and most probably is assessing, to some degree, two or more of the subject traits.

TEST DESCRIPTIONS

The following provides a description of the above mentioned tests.

1. The perceptual psychomotor test consists of on two solid state perceptual psychomotor tests based in part on the two-hand coordination and complex coordination tasks (stick and rudder test) developed during World War II. Sanders, Valentine and McGrevy (5) report that both tests were transfigured into a solid-state indepdndent testing apparatus of high reliability. Subsequent validation of the tests indicated that complex coordination was a reliable and valid predictor of success vs. failure and flight training deficiency (similar to the Navy term Flight Failure) in Undergraduate Pilot Training (UPT). McGrevy and Valentine (3) report that the perceptual psychomotor complex coordination test made a unique contribution to the prediction of graduation from Air Force UPT above and beyond that provided by the Air Force paper-and-pencil test selection instrument, the AFOQT.
2. The Integrated Multi-Task Psychomotor and Cognitive Testing (IMPACT) system assess the performance characteristics of poetntial aviation personnel heretofore not evaluated by standard assessment techniques. As opposed to traditional, system-specific flight simulators, the tasks comprising IMPACT are synthetic and are designed to assess the information-handling capabilities and capacities of the human operator. Human attentional capacity and perceptual-motor capabilities have already been separately shown to relate to flight criteria such as attrition from training programs. IMPACT, however, is an integrated tasking system which simultaneously assesses those information processing behaviors (i.e., perceptual-motor, attentional, time-sharing) most critically related to performance in actual aviation systems. Tasks employed by IMPACT are easy to learn and are fully automated to reduce test administration time. Moreover, the modular nature of these tasks provides the inherent flexibility required to readily present test batteries for Naval Flight Officer testing as well as naval pilot testing. The measurement methodolog employed will circumvent traditional problems associated with assessment of performance in multiple-task, multiple-measurement situations. The use of adaptive task loading techniques during single task learning acquisition phases will be used to adjust the total tasking configuration to an individual's own skill level. These methods have been demonstrated to provide excellent measurement parsimony and high measurement reliability (4). The objective of IMPACT is to use these techniques for identifying the potential capability of aviation personnel and to relate these assessments to operational flight performance criteria.
3. The dichotic listening test investigated by the Israeli Air Force was successful in predicting aviator training success in high-performance jet aircraft. In dichotic listening assessments, messages are presented to both ears simultaneously. The subject is told to ignore the message on one auditory channel while listening for relevant words or items on the other channel. Gopher and Kahneman (2) report validities in the range .30 to .40 using 100 Israeli flight candidates. More recently, Gopher (2) presented new data on the dichotic listening test based on a population of 2000 individuals during training. His initial findings indicated a low but significant relationship with success in jet training ($r = .18$). Although the correlation was low, it has virtually no relationship with other predictor measures, offering a potential new dimension to the prediction of pilot success.
4. The Brief Vestibular Disorientation Test (BVDT) involves observer assessment of subject's reactions produced by head movements in a rotating chair. The reliability of observers' assessments has been demonstrated and significant validation and cross-validation coefficients have been reported for criteria of pass versus various types of separation from training. It has also been established that the BVDT score significantly augmented the multiple correlation of existing aviation selection variables with the same criteria (1).
5. The Automated Pilot and Aptitude Measurement System (APAMS) was originally developed using general aviation trainers (GAT) and flight scenarios. Those scenarios developed and validated by the Air Force have been obtained and will be modified as required to meet Navy testing needs. Rather than use a GAT type device, it is planned to utilize a device developed by the Naval Training Equipment Center referred to as the TRAINALATOR. It is a device that can be programmed to provide the flight characteristics of various aircraft for testing cognitive and psychomotor skills or refreshing previously learned flight skills. The device consists of a digital computer to perform operations concerned with the management of training, control of instructional scenarios, scoring, simulated aircraft characteristics, and to drive the instructional display. Instruction is provided by a voice generation system and a CRT. Student inputs are made through a small "joystick" and throttle. Simulated aircraft instruments with indications in real time are displayed on the CRT. A facsimile of the T-34C stick and power level could replace the "joystick" and throttle. A simple cockpit environment could be provided as desired. The device has been demonstrated as a higher order, partially self-organizing flight training system with testing functions.

OPERATIONAL MODEL

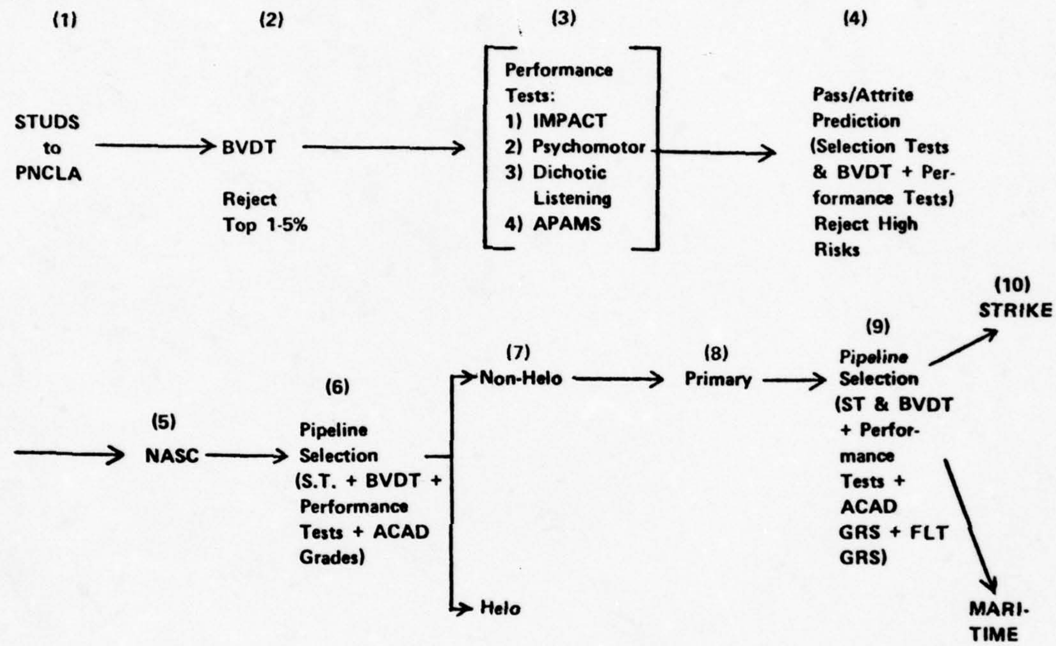
Operationally, the system might well take a form outlined in Figure 2. The students will report to Pensacola for their final screening prior to being accepted into the program. This final screening will be for a one week duration and will include a thorough physical examination. FAR and AQT scores will have already been collected. They will then be assessed on the BVDT. Highly susceptible applicants will be rejected out of hand. Research has shown that the top 2%-3% have very little chance of completing the program. They will then be assessed on the dichotic listening tasks, the psychomotor task, IMPACT, and APAMS. At this time, a multiple regression equation will be applied using all the information available up to that point. Those applicants with high risk scores would be rejected, while the remaining would be allowed into the program. After completing the Pre-Flight phase of training, another multiple regression equation would be applied with certain training grades now being added. The result would provide an input into deciding (should it be necessary) who should and should not be assigned to Helos. A reiteration of the procedure would again be applied after the student has completed Primary Flight Training. This time the data would be used in assisting in the Strike versus Maritime pipeline decision.

CONCLUSION

The selection system described above attempts to include an efficient mix of predictive power, administrative facility, and cost effectiveness. Final estimates for the efficiency of the mix must await the conclusion of the 18-month pass-through study. However, the author of this paper, at least, feels confident that a substantial improvement in selection decisions for naval aviation training can result, and, as anyone familiar with the costs involved in aviation training can attest imporvements in selection decisions quickly translate into significant cost savings.

FIGURE 2

Proposed Selection System Model



REFERENCES

1. Ambler, R.K., and Guedry, F.E., Reliability,and validity of the Brief Vestibular Disorientation test compared under 10-rpm and 15-rpm conditions. *Aerospace Med., 42*: 186-189, 1971.

2. Gopher, D., & Kahneman, D. Individual differences in attention and the prediction of flight criteria. *Perceptual and Motor Skills, 33*: 1335-1342, 1971.

3. McGrevy, D.G., Valentine, L.D. Validating two aircrew psychomotor tests, AFHRL-TR-74-4. Personnel Research Division, Lackland AFB, Texas, 1974.

4. North, R.A., Harris, S.D., and Owens, J.M. Test-retest reliability of individual differences in dual-task performance. NAMRL TR 1-78. Pensacola, Fla., 32508.

5. Sanders, J.H., Valentine, L.D., and McGrevy, D.F. The development of equipment for psychomotor assessment. AFHRL-TR-71-10. Lackland AFB, Texas, Personnel Division, Lackland AFB, Texas, 1971.

THE USAF OCCUPATIONAL SURVEY PROGRAM – AN AID TO FORCE MANAGEMENT[1]

James A. Turner, Jr.
USAF Air Force Occupational Measurement Center
Lackland AFB TX

As managers in today's Air Force we are constantly looking for better ways to lead the force. Yet it is surprising how often key decisions are made without specific, objective data. I am referring primarily to people decisions since accurate data about hardware seems to be more readily available. With people, however, we managers frequently don't have the information we need or, more typically, assume we have good information because of our own personal experience. In the latter case this amounts to no more than a high level of "seat-of-the-pants" management because our experiences are just that - personal, situationally specific, and often limited.

I want to talk to you today about a source of objective data that is available and discuss how it can be of use to managers. I am referring to occupational survey data, data that is the outcome of the USAF Occupational Survey Program. This program is a function of the USAF Occupational Measurement Center, an Air Force aagency with an AFMPC OPR but located within ATC. We produce a fine product and I feel it important that our capability be known to the Air Force specifically and DOD in general.

First, a few words about our background. The USAFOMC, as it is known today, grew out of the Air Force Human Resources Laboratory (AFHRL). AFHRL was engaged in developing and testing occupational analysis methodology since the late 1950s. In 1965 the program was ready to go operational and in July, 1967, ATC began the operational survey program. After a series of expansions the USAFOMC was created in 1974. Consequently, even though our history as a Center is brief, our roots and the validity of our procedure extend back for over 20 years. Even in our brief existence as a separate agency we have still surveyed over 250 officer and enlisted career areas and over a half a million jobs.

Now a brief introduction to the occupational survey program. We start with a task list development phase in which task writers totally immerse themselves into a particular Air Force Specialty. The primary objective of these developers is to create a behavioral task list that encompasses the entire career field. This is quite an art, for the task list developer usually has only a layperson's understanding of the field. Nonetheless, at the end of 12 weeks a task list, based on reference materials and interviews with subject matter specialists, has been written. The task inventory is then validated by operational personnel to insure its completeness. This point should be emphasized here - the task list is a function of career field personnel and not simply a product of someone's personal ideas.

Concurrent with the development of a task list, we also create a detailed background section. In the background section we ask a series of questions aimed at the respondent's age, sex, grade, command, duty title, equipment worked on, job progression, shift worked, job satisfaction, and in general any other information that might be of interest. These background items are extremely critical for they allow us to later categorize people on any dimension we wish. The final survey booklet, or inventory, includes both the task list and the background section.

This inventory is then administered to the career field. We survey 100 percent up to a manning level of 3,000; beyond that we use a stratified technique. Respondents tell us what tasks they do in their *present* job and how much time they spend, relatively speaking, doing them. Once the survey booklets are returned, they are processed by the AFHRL computer using a complex series of statistical programs created especially for the survey program. At this point, an occupational analyst takes over and in a period of 8 weeks analyzes the results in terms of paygrade, AFSC, job structure, and any other factors pertinent to that specialty. The end result is an Occupational Survey Report (OSR) which is distributed to training, classification, and user personnel.

We analyze 51 enlisted career areas a year; that works out to surveying each specialty every four years. In addition, we are engaged in a wide variety of special projects. One such project, for instance, is the Electronic Principles Inventory, which was discussed at this conference in 1976. Last year we provided specific feedback to training managers about the use of electronic principles in 54 career fields. Last year we also completed a special project that provided specific feedback to Tacticel Air Command about how their Fighter Weapons Instructors (cost to train: $225,000+) are being used.

We are conducting an officer survey program. Last year we produced an OSR for the Security Police Officer career area; this year we will produce OSRs for the Weapons Controllers (17XX) and the Space Systems Analyst (20XX) career areas. Presently we are engaged in other projects for Air University, the Defense Intelligence Schcool, the Federal Procurement Institute, and the AWACs program. Many of these special projects and officer surveys are leading us into the executive-leadership-management areas. We welcome this tasking for it is apparent, to us at least, that the occupational survey program can provide accurate data about what managers do. This new data will have quite an impact as doubtlessly will be the case when we finish our Professional Military Education Project for Air University.

That's great you say. I think I understand what you do, but how will it help me? Logical question, for it is difficult to immediately see how data about percent members performing a task and relative time spent can be useful. But look at an item as simple as how your specialty is described in AFR 39-1 or 36-1. The majority of our reports recommend changes to the specialty description. An accurate job description would appear to be a must for force managers. It also assists a recruiter immensely. Its very informative to personnel wishing to transfer into the specialty. On the other hand it is a little bit surprising to discover that inaccurate job descriptions exist. How can proper force management decisions be made? Wasn't there someone who realized the inadequacy? Shouldn't a manager or trainer be expected to know the entire job spectrum? We think so, and our data provides that information.

Or look at the Specialty Training Standard (STS). The STS impacts, indirectly, on the Specialty Knowledge Test (SKT). What happens if the STS is not accurate? Simple, you get an SKT that is not a true reflection of what is being done in the field. Result - frustration on the part of the enlisted troop taking the SKT. Morale problem - you bet. Our data can tell you how accurate your STS is. It can do the same for a course Plan of Instruction or POI; essential information for a manager. Simply stated, our data can tell you what is being done in the field. You can compare that information against any document you want. Too often we have seen education programs assessing what someone thinks is happening versus what is happening; that should never be the case.

[1]The views expressed in this paper represent those of the authors and do not necessarily reflect the views of the United States Air Force or the Department of Defense.

Our data can also breakout job performance by skill level, paygrade, and military service. It can tell you what is being done by senior, but not by junior, personnel. It can tell you when certain tasks (such as supervisory tasks) start. Ever wondered about assignment specificity? We can tell you. We can break the data down to shop level and lower. If you think such information would be useful in job design, you are right. *I could test that statement right now. How many of you in this audience know how your subordinates spend their time?* We do. Enough said.

Such issues have been the normal questions answered by our data. Recently we have begun to see other ways to assist managers, not by any great changes in our methodology but by simply getting people to understand what our data can do. For instance, with our first-term job description we have been able to sit down with MAJCOM users and say, "Here is what the new guy does in his first assignment. Which of these tasks do you want him trained for when he hits your base and which would you rather train him for via OJT?" Once we have this agreement, the school has a much clearer definition of what they are obligated to train. Correspondingly, MAJCOMs perhaps for the first time understand how their high turnover personnel can be best utilized.

We have also recently realized how our data impact on command policy. It has not been uncommon for a report to state that, for instance, *AFSC 10170 personnel are performing a certain job. When our report hits the MAJCOM, we sometimes hear, "No, 10170s don't do* that." At that point we promptly step aside; we are not implementers. It is not our decision to say whether actual utilization or command policy should change. We are simply providing management the opportunity to see what is happening. If all is okay, thats fine with us. If not, our data provides a base from which to act. The critical point here, however, it that we can identify such a discrepancy. Our data can tell you when what is assumed is not what is.

We have also begun to realize how our product impacts in ways we never considered. Our report, once published, becomes the frame of reference for all future decisions about that career field. Like it or not, the objective analysis we provide is always going to be considered. Suggestions that go against our report have to be justified much more than those in line with our data. Also, our report won't just fade away. We recently had a situation where the project requester did not take action on the results we provided. If became obvious to us that it was a case of "let's don't do anything and it will go away." However, that agency's auditor obtained a copy of our report, and now all kinds of questions are being asked. Again, we were not directly implementing but our data did lead to changes, nonetheless.

This note may be a good place to end. We are objective data providers. We have no ax to grind. We just want to take a snapshot of your field, analyze what we see, and give you the results. We do that, and we do it well. That's why the Occupational Survey Program is an aid, an extremely valuable aid, to managers.

DIFFERENTIAL PERCEPTIONS OF AIR FORCE JOBS: CONFLICTING PERSPECTIVES

By

Walter E. Driskill, Ph.D.
James B. Keeth
Jimmy L. Mitchell
Occupational Survey Branch
USAF Occupational Measurement Center
Lackland Air Force Base, Texas 78236

One major cause of malaise concerning the Air Force personnel system lies in the differential perceptions about what jobs are or should be. This is a problem in most personnel systems throughout government and industry and is an area which is not yet well studied or completely understood. There is some research, however, which suggests that the disparity can be very great between how workers and supervisors perceive the same jobs.

O'Reilly (1973) studied 57 pairs of supervisors and workers using task lists developed through work diaries. He found that in only one of the 57 cases was there full agreement between the supervisor and worker on what tasks were actually required. Fifty-two of the 57 supervisors reported workers as doing tasks which the workers said they did not do and four-fifths of the supervisors said that certain tasks need not be performed which workers said were required. Thus, we can see that even in the fairly simple situation of first-line supervisors and their workers, there is often quite different perceptions about the content and the skill requirements of jobs.

When we deal with multi-level personnel structures, such as the military or any large industry, the situation becomes even more complex. There has been some interesting research which suggests that level of organization is a key variable in terms of the perception of performance or behavior. Bernardin and Alvares (1975) studied such perceptions and concluded that organizational level was one of the two significant variables involved and that there were also significant interactions among variables. For our puposes here, it is sufficient to recognize that this research confirms that perceptions of the individual's job and performance do vary according to the organizational level of the observer.

For the military situation, with its many different levels of organizational structure, the implication is that each of these levels will have differential perceptions as to what tasks are involved in a given job. Indeed, we find this implication confirmed in actual exprience; depending upon their level of assignment, managers and staff personnel perceive jobs differently, and consequently have varying expectations about the tasks involved, the skill requirements of the workers, and even of the quality of performance of workers who fill the jobs. Even more critical, however, is the fact that these perceptions often match neither the real world of job structure nor the classification, training, assignment, and career progression policies of the formal Air Force personnel system.

The differential perceptions of jobs seem to vary from the very macro level to the very micro. The functional manager of a career area operating at the Air Staff level tends to be the most macro; here the expectation is of a worker who is trained and experienced in all aspects of his Air Force specialty. Function managers seek he ideal Airman who can go anywhere and do anything involved in the occupational field, but of course this species turns out to be very rare indeed.

Major air command staff personnel tend to have a slightly more restricted view; their concern is not for someone who can do every task involved in a career field but only those tasks involved in their own command. At Base level, the expectations are even more limited but still require that the individual be qualified to work on any system on the base. At the Unit level, commanders want an individual who can perform in any section, while immediate supervisors are mainly concerned that the individual be able to do any task on a specific job. As was discussed earlier, however, the supervisor's perception of what constitutes the job may be quite different from the perception of the individual worker.

Expectations about the quality of experience of the technician obviously differ among these various organizational levels. Such expectations are further complicated by variations in the nature of work structure among different Air Force specialties; and in many cases these variations in actual work structure create problems, since they may not be consistent with the general Air Force personnel policies. The generalized training, assignment, career development, and career progression policies of the Air Force personnel system work well in some types of specialties but may actually inhibit the development and progression of individuals in other types of occupations.

A key question is to find or establish some objective criterion of what an Air Force specialty is and should be. O'Reilly in his study of the differences in job perceptions between supervisors and incumbents concluded that it was difficult to determine which description of a job was the correct one. He recommended that this was a problem which would have to be resolved through job analysis (O'Reilly 1973).

The Air Force was a pioneeer in developing an operational job analysis system to assist in providing a better definition of what Air Force spcialties are and should be. The Air Force occupational analysis program has been operational for almost 11 years (Driskill 1977) and in this time, over 600,000 personnel in 250 Air Force specialties have been studied. This process requires detailed descriptions of a specialty at the task level by technicians who are highly qualified in the specialty; administration of this structured task list to very large samples of skill levels at all the locations of the specialty; and analysis of the resulting data through the Comprehensive Occupational Data Analysis Programs (CODAP). This process has been detailed elsewhere (see Morsh 1964; Morsh and Archer 1967; Christal 1974; Driskill 1977; or Keeth 1977); it is sufficient to recognize that this is a well established and accepted job analysis system which is now used by all the US military services, by industry, and by the civilian academic and training community as well (McCormick 1976; Ammerman 1977).

Occupational analysis provides many different kinds of basic information about an Air Force specialty. First, it provides information about the kinds and numbers of personnel who perform each task. Second, it can provide a job description for any selected group of individuals — by skill level, time in service, major command, duty location, or any other pertinent variable. Third, a CODAP analysis can also determine how the work in a speciality is actually organized; that is, how many different kinds of jobs are being performed and how these jobs are related to one another. The jobs performed by individual workers can be grouped on the basis of the similarity of tasks they perform. Such a grouping leads to the identification of the specialty.

As a result of looking at the 250 specialties over the years, we can also begin to see some general patterns and problems. Three highly significant conclusions which emerge from our years of occupational analysis experience are that:

1. No single worker ever performs all the tasks of the specialty in a single assignment.

2. Few, if any, individuals ever perform *all* the tasks of the specialty even over an entire career.

3. Specialties vary quite drastically in the degree of specialization of individual workers; some specialties consist of only a few types of jobs where others (and probably most) consist of a number of discrete.jobs.

Those specialties consisting of only a few jobs are considered homogeneous specialties. Quite often a homogeneous specialty will consist of one major job cluster, where everyone is doing basically the same job, along with several very small or highly specialized job types (typically including some supervisory positions, technical school instructors and perhaps technicians performing staff functions at major command level or higher). The Aircraft Loadmaster (AFX 114XO) is a good example of a homogeneous career field. Most Loadmasters perform essentially one job—that of loading and unloading aircraft. Job incumbents perform similar sets of tasks regardless of whether they work on a C-5, C-141, or C-130 aircraft, or whether they are a 3-,5-, or 7-skill level specialist or technician. Once certified as a Loadmaster, they perform basically the same job throughout their career, although obviously the more senior Loadmasters eventually perform supervisory, training, and staff functions in addition to the basic Loadmaster job. This type of homogeneous career field facilitates haveing a single common training program (with later specialization by aircraft system), a well-structured career development course, and subsequent promotion testing. In fact, a recent review of this area by representatives of MAC, the Loadmaster school personnel at Sheppard, and OMC resulted in only minor changes in the official specialty description (AFR 39-1) and training standard (STS). The only changes in training (AFR 39-1) and training (STS). The only changes in training recommended were those aimed at avoiding duplication between the basic (ABR) technical training course and the aircraft-specific MAC courses which are designed to certify aircrew status on C-5, C-141, or the C-130.

Specialties consisting of a large number of diverse and discrete jobs are at the other extreme of the homogeneity-heterogeniety continuum. Heterogeneous specialties consistently show job incumbents working only in small specialized areas within the overall specialty. Typically, such jobs are quite discrete and consist of very few tasks in common. Incumbents may be required to work in any one of many different and often quite unrelated jobs. A good example of such heterogeneous specialties would be the Medical Administration career field. In our analysis, we identified 60 different jobs being performed; many of these job types were unrelated to each other in terms of the tasks performed. Another example would be the Supply Services area which includes Commissary workers, Billeting Clerks, Mortuary Affairs NCOs, Linen Exchange, and Clothing Issue personnel. You can tell just from these job titles that these are discrete, nonoverlaping jobs and that any type of common training for these individuals would be impractical and uneconomical.

Such heterogeneous specialties represent a very difficult challenge in terms of developing training programs, designing career development packages, and constructing promotion tests. It is difficult to provide common resident training programs for this type of specialty; in the medical administrative area, such training would have to cover 60 different types of tasks. Yet to provide such general training would not be effective, since the individual would probably perform at most only one or two types of jobs during his initial enlistment. Thus, any training beyond that covering those two types of tasks would represent substantial overtraining which would never be used by the majority of medical administration technicians.

In many heterogeneous career fields, it is possible to develop effective training programs by channelizing training, that is, by providing that training which the individual needs in order to perform effectively on his first job. This is now the Air Force policy whenever possible and is rapidly being implemented. Channelization is done by establishing formal shredouts (subspecialties); by limiting training to only one job type (where there is one major cluster); or by finding some other way to identify the first assignment of the individual in time to properly tailor training (as was recently done in aircraft maintenance where it is now possible to specify the major aircraft system). However, where there are too many job groups or where the personnel system cannot identify the first assignment of the individual, then it is not possible to provide effective resident training programs for highly heterogeneous career areas. It may be that for these specialties, on-the-job training is the only workable solution, as is the case with Supply Services. However, even with effective OJT programs, the problems of developing a comprehensive career development course (CDCs) and constructing realistic and effective Specialty Knowledge Tests (SKTs) remain.

In the heterogeneous career fields, the problems of OJT become almost insurmountable. The normal Air Force assignment system assumes that an individual in a specialty can perform any job within that speciality, yet in a heterogeneous career field, the person may have neither the training nor the experience to do the new job. Thus, In such fields, the individual must undergo extensive OJT with almost every new assignment. This is disadvantageous to both the individual and to the Air Force, since it requires an inordinate investment of the individual's and the OJT trainer's time, thus reducing the amount of productive work which can be accomplished.

These findings imply that the resident training program, formal on-the-job training, career development course, promotion tests, assignment system, and even the classification structure needs to be tailored specifically to the needs of the particular specialty, based in part on the relative heterogeniety of the specific occupational field. We need to use whatever objective information we can develop, such as that developed in the Air Force occupational survey program in conjunction with the expertise of career field specialists (representing all commands using the specialty and preferably representing several levels of the specialty within the command from individual units to command headquarters) to examine how incumbents are actually being utilized. From this data base, it is possible to determine how such career field specialists *should* be used and, in conjunction with trainers, OJT specialists, CDC writers, SKT developers, assignments personnel, classification monitors, and functional managers at the Air Staff level, to design more effective career assignment, training, career development and promotion programs.

It is extremely important that all of the diverse perspectives be represented in any decision involving the personnel programs relating to an occupational field. Such programs are currently so interrelated that a change in one program often has drastic effects on the other programs, sometimes with very unexpected results. Only by sitting down and discussing the entire career field and its associated programs can such problems be anticipated and resolved.

We feel that in our current occupational survey program we can provide a common data base which can be the starting point for such discussions and which can be used to validate the existence of various potential problem areas. Such data provide a common catalogue of the tasks involved in a specialty which, since the task statements were developed by career field incumbents, are in the language of

the specialty. Since such data provide a common language, they are exceptionally useful as the initial focus of utilization conferences since they can be used to define the various positions (perceptions) resulting from differing organizational levels or purposes represented. Once such differences are defined in commonly understood task-oriented terms, they can be more easily understood and modified. As Drucker has pointed out, "to realize that they see the same reality differently is in itself already true communication" (Drucker 1970 as cited in O'Reilly 1973). Once such communication is established, then many of the differences in perceptions of Air Force jobs can be resolved and many of the specialty-specific problems can be solved.

In the Air Force occupational survey program, we find that we have a rather unique perspective ourselves. We are in the unusual position of having considerable interaction with many of the organizational levels involved in most career field; from the unit level where we work in developing task lists all the way up to the Air Staff functional manager, where we discuss our final Occupational Survey Report (OSR). By interacting with these various levels, we are continually learning both about the specific career field and about the possible uses of our occupational survey data.

Further we find that our perspective of a career field changes over time as we resurvey career fields for a second and third time. Where our initial objective was to review the classification structure and our second was perhaps to examine the resident training programs in some detail, by the third survey we may find ourselves more interested in the career development program, the field OJT program, or promotion testing. In each case, we find new ways to use occupational survey data to assist in solving real-world Air Force problems. This often requires that we develop new analysis programs in conjunction with the Air Force Human Resources Laboratory or that we find new ways to display data from existing programs.

We also find that as our interaction with all of the units and agencies involved with a specific career field grows, so does our perspective of each career field as a unique and complex system. We are increasingly conscious of the need to examine each specialty in detail and to understand all of the programs and problems associated with it. We know that each field varies in terms of its tasks, its demands on the individual, its training, etc. Thus, we are increasingly convinced that a separate "model" is necessary for each career area and that we should not expect broad Air Force programs (such as technical training, WAPS testing, CDCs, STS, etc.) to be equally applicable or valid for every specialty. Separate models may also be required in the officer utilization fields as well, but our experience with officer studies is as yet somewhat limited.

The point that we wish to emphasize here is that it is essential that we develop more coherent models for each Air Force specialty. Included in such a model must be some identification of the various organizational levels which are associated with the specialty and thus a specification of the different perspectives which can be expected. Once these various potential perspectives are identified, they can then be contacted and in the resulting interaction, the differences in perspective can be clarified. This is perhaps the critical initial step which must be taken if we are eventually to work together in identifying and resolving the problems of a specialty.

BIBLIOGRAPHY

Ammerman, Harry L., Relating task surveys to the content of existing training programs. Paper presented at the Military Testing Association Meeting, San Antonio, Texas, October 1977.

Bernadin, H. J., and Alvares, K. M., The effects of organizational level on perceptions of role conflict resolution strategy. *Organizational Behavior and Human Performance*, 1975, *14*, 1-9.

Christal, Raymond E., The United States Air Force occupational research project. Lackland AFB, TX: Occupational Research Division, AFHRL, AFHRL-TR-73-75, January 1974.

Driskill, Walter E., The USAF occupational analysis program: a symposium. Lackland AFB, TX: USAF Occupational Measurement Center, Technical Note 77-01, October 1977. (Papers presented to the Military Psychology Division at the 85th Annual Convention of the American Psychological Association, San Francisco, August 1977).

Keeth, James B., The USAF occupational analysis program. Paper presented at the Military Testing Association Meeting, San Antonio, Texas, October 1977.

McCormick, Ernest J., Job and task analysis. in Dunnette, Marvin D. (ed) *The Handbook of Industrial and Organizational Psychology*. Chicago: Rand McNally College Publishing Company. 1976 Chapter 15: 651-696.

Morsh, J. E., Job analysis in the United States Air Force. *Personnel Psychology* 1964, 17: 7-17.

Morsh, J. E., and Archer, W. B., Procedural guide for conducting occupational surveys in the United States Air Force. Lackland AFB, TX: Personnel Research Laboratory, AMD, PRL-TR-67-11, September 1967.

O'Reilly, A. P., Skill requirements: supervisor-subordinate conflict. *Personnel Psychology*, 1973, *26* 75-80

OCCUPATIONAL ANALYSIS OF OFFICER UTILIZATION FIELDS

Rodger D. Ballentine
Paul N. Ditullio
David S. Street
and
Captain David S. Street
Occupational Survey Branch
USAF Occupational Measurement Center
Lackland AFB, TX 78236

A well established occupational survey methodology has been developed based on twenty years of research by the Air Force Human Resources Laboratory (AFHRL). Survey Procedures evolved after job inventories were constructed and administered for 85 airmen career ladders and 16 officer utilization fields during 1961-1967 (Morsh & Archer, 1967). An additional seven officer utilization fields were surveyed between 1971 and 1972 using the job inventory approach. Over the past 11 years these survey procedures have been used in job analyses of over 250 career ladders by the USAF Occupational Measurement Center (USAFOMC). Since 1974, the USAFOMC has operationally studied several officer utilization fields.

For the purpose of his paper, we will assume you are familiar with the traditional occupational survey process. More information about the survey process is provided in Morsh and Archer's procedural guide (1967). In 1975, Mayo, Nance, and Shigehawa evaluated the use of the traditional job inventory approach for analyzing USAF officer utilization fields. The basic finding of this report was that the inventory method can be used operationally in the analysis of officer jobs. These authors indicated that to effectively study officer utilization fields some modifications of traditional job inventory development procedures are necessary. Changes specifically suggested were:

1. A more detailed examination of the utilization field to be studied is required prior to task list development. This examination should include gaining information about organizational structure, review of local job descriptions, and an overview of general, rather than task specific, work performed.

2. Developing officer job inventories necessitates conducting significantly more interviews and field reviews to develop the task list. "It is not unusual for thirty interviews and one hundred field reviews to be required for inventory construction" (Mayo, Nance, & Shigekawa, 1975, p. 82).

3. A detailed task list rather than one constructed of broad general statements results in better discrimination of jobs within a utilization field. The optimal level of task specificity must be determined for each field studied.

As noted earlier, the USAFOMC has, since 1974, conducted occupational surveys of officer utilization fields. Surveys have been completed for Security Police Officers, Communications Electronics Officers, Signals Intelligence Officers, and a subgroup of tactical fighter pilots. In addition, job inventories have been developed for Space Systems Operators and Analysts, Weapons Controllers, and Procurement Officers. These initial attempts to operationally survey officer fields with traditional methodology have highlighted procedural differences similar to those outlined by Mayo, Nance, and Shigekawa (1975).

In order to cope with the procedural differences noted, the occupational analysts at the USAFOMC employ the following techniques:

1. Prior to the development of an inventory for an officer utilization field it is necessary to become thoroughly familiar with the variety of jobs performed, the training provided, the career management of incumbents, and the projected uses of the occupational data. Typically, the best place to begin is with career managers at the Air Force Military Personnel Center. Interviews with appropriate training and functional managers are also helpful in defining the diversity of jobs performed in a utilization.field. This thorough front-end analysis of the officer utilization field structure is necessary to determine the factors which differentiate the jobs performed, such as location, major command, unit or mission. As was noted in the survey of tactical fighter pilots (Ballentine, 1977) a prior knowledge of the aircraft missions was an important factor in differentiating the tasks performed by the pilots surveyed. Only when there is a clear understanding of the utilization field structure can the job analyst develop an effective survey strategy. This information is also necessary to determine the appropriate survey sample and to identify key field personnel to be interviewed for task list development.

2. Interviewing a representative sample of officers performing the various jobs in an officer utilization field is difficult. The fact that few incumbents perform the same job and the limited time officers have to be interviewed make it difficult to arrange necessary interviews. Also, there are fewer officers in any given job at one location, which limits the opportunity to hold the group discussions typically necessary to identify different task functions or to obtain agreement about the wording of task statements. In order to write generally understandable task statements, it has been necessary to bring together officers performing the same job for task list development interviews. Group meetings of officers with different jobs have also been held to validate the task list. These conferences provide an opportunity for officers with various jobs within an officer specialty to discuss the wording of tasks, check redundancy, and eliminate overlapping tasks. Validation conferences have been effectively used in developing several task lists.

3. Analyses of officer survey data are complicated because of the aforementioned difficulties in developing task lists with clear detailed tasks. Normal procedures used to discriminate among officer task performance may be negated by the large number of broadly written, general supervisory and managerial tasks in officer surveys. Such tasks are usually performed to some extent in most officer supervisory jobs. When this occurs, frequently the analyst is unable to delineate differences in the jobs performed. This reduction in the effectiveness in differentiating among jobs performed by officers may be avoided by more detailed task lists. For example in the occupational survey for Signals Intelligence Officers (DiTullio, 1976) members in all jobs spent 50 to 80 percent of their time on supervisory and managerial tasks; however, there were enough specific technical tasks that jobs could be differentiated.

Overall, officer job analyses are more complicated and require more time, money, and personnel to accomplish. If occupational analyses of officer utilization fields are to provide meaningful information to personnel and training managers, existing survey procedures must continue to be adjusted as described. The survey development process must be planned to deal with the difficulties noted during officer task writing interviews. An effort must be made to bring together officers from diverse jobs who can be interviewed simultaneously for an extended period of time. Every effort must be made to write task statements which adequately describe the possible tasks performed by officers to be surveyed. Finally, analysis of officer survey data must be done with an awareness of how commonly performed managerial tasks may disguise more subtle job differences. Recognition of the necessity for procedural differences is critical to further job analysis of officer utilization fields.

BIBLIOGRAPHY

Ballentine, R. D., *Occupational Survey Report of Fighter Weapons Instructors* (AFPT 90-SXX-333). Lackland AFB, TX: Occupational Survey Branch, USAF Occupational Measurement Center, October 1977.

DiTullio, P. N., *Occupational Survey Report of Signal Intelligence Officer* (AFPT) 90-803-227). Lackland AFB, TX: Occupational Survey Branch, USAF Occupational Measurement Center, August 1976.

Mayo, C. C., Nance, D. M., & Shigekawa, L., *Evaluation of the Job Inventory Approach in Analyzing USAF Officer Utilization Fields*. AFHRL-TR-75-22. Lackland AF, TX: Occupational and Manpower Research Division, Air Force Human Resources Laboratory, June 1975.

Morsh, J. E. & Archer, W .B. *Procedural Guide for Conducting Occupational Surveys in the United States Air Force*. PRL-TR-67-11, AD-664 036. Lackland AFB, TX: Personnel Research Laboratory, Aerospace Medical Division, September 1967.

THE STUDY OF EXECUTIVE & MANAGERIAL JOBS

Jimmy L. Mitchell
Occupational Survey Branch
USAF Occupational Measurement Center
Lackland Air Force Base, Texas 78236

There is a growing interest in the study of executive and managerial jobs, both in the military community and in civilian universities and industry. Most of the job analysis systems which have been developed over the last four decades of occupational research have attempted to analyze and evaluate all types of jobs and the sample populations studied have most frequently been blue collar occupations. In most companies, such blue collar positions far outnumber the executive, professional, and managerial positions. As a result, most occupational studies have been heavily weighted in favor of production and line workers.

In the military services, the same type of emphasis has been evident. The vast majority of occupation analysis studies have focused on enlisted jobs since the bulk of our resident training courses, on-the-job training, career development courses, and other programs have focused on the enlisted force. However, in the military there have been a number of studies which have dealt with officer career areas where there is an unusually high training investment cost, such as rated officers, medical personnel, and research scientists. Thus, military occupational analysts have been aware of the need for separate officer studies but have not given them high priority except where they represent a high value resource.

One major difference between civilian and military job analysis programs has been in the basic approach. Military studies have generally been done within specific occupational areas and have been oriented to examining the tasks performed (Morsh 1963). Civilian job analysis systems, on the other hand, have generally been across occupational areas and have focused on the common factors or dimensions which underly jobs (Fine 1967, Baehr 1967, McCormick, Jeannerret, & Mecham 1972). These differing approaches result in part from the different uses of job data in the military and civilian communities—military job analysis generally focuses on job definition, classification structure, and specification of training requirements where civilian job analysis systems most frequently involve identification of employment opportunities and applicant referral (Departmnt of Labor 1972),determining salary levels or job evaluation (Scott 1963), and improving our scientific understanding of the world of work (McCormick 1976).

Increasingly, however, the two approaches to job analysis; that is the task-oriented military approach provened by Morsh & Christal (Morsh 1963) and the worker-oriented civilians approach (Fine 1963; McCormic, Jeannerret, & Mecham 1972) have been cross utilized. The Air Force task analysis system has been adapted by several groups for the study of Vocational & Technical curricula (Ammerman 1977). In turn, the civilian job evaluation regression system has been adapted to the determination of appropriate grade levels for military jobs (McCormick & Mecham 1969; Christal 1965, 1975).

During the last several decades, there has also been a growing awarness that job analysis techniques developed for the study of blue collar jobs may not be completely appropriate for the study of executive, management, and professional positions. I would like to very briefly review some of the studies which have been done on management-level positions and and to suggest some of the directions we seem to be going in our understanding of such higher level jobs.

In a pioneering study in the late 1940s, Colonel John Flanagan studied Air Force Lieutenants and Colonels in terms of their activities and responsibilities (Flanagan 1951). His work impacted the Officer Effectiveness Rating system for a number of years and let him validate the critical incidents technique for the analysis of management jobs.

Fleishman (1953) studied leadership by factoring data on a variety of leader activities and characteristics. His resulting production and consideration factors have greatly influenced the study of leadership and resulted in whole schools of leadership theory and training programs.

Williams (1956) studied executives using the critical incidents approach and developed a set of dimensions very similar to those of Flanagan.

Shartle & Stogdile (1957) pioneered in the use of time allocation across a managers responsibilities and activities. This approach was not generally accepted for civilian use but had a considerable impact on military job analysis.

Hemphill (1960) developed the first major job analysis instrument designed specifically for executives and his work has become a classic. His Executive Position Description Questionnaire (EPDQ) resulted in the identification of ten factors underlying executive performance. Unfortunately, follow on research by others generally failed to replicate his factor structure (Pruen & Ronan 1970) and a number of questions concerning his methodology have been raised. His factors appear to be a mixture of activities, personal characteristics and social conditions which are difficult to interpret and use in specific job analysis applications. Most of his factors have recently been replicated in an extensive study by Tornow & Pinto (1976).

Mahoney, Jerdu,& Corral (1963, 1964) examined managerial performance across a set of eight types of specific activities. Their work is noteworthy in terms of their use of a consistent set of activities which could be realistically rated by a single rating scale.

Baehr (1967) desired a set of 12 managerial factors using a structured questionnaire, a single rating scale, and a large sample of managers and executives. Her work was instrumental in validating the use of a structured questionnaire specifically designed for managers.

McCormick, Jeanneret & Mecham (1969, 1962) developed the Position Analysis Questionnaire (PAQ) for the study of all types of jobs. This system has become one of the most widely used job analysis systems. It has been used in the study of executive positions in the US Navy (McCormick & Mecham 1969) and has demonstrated considerable potential for use in job evaluation, selection screening, test validation, and a variety of other uses.

However, the question of the relevancy of all PAQ items for the special subset of executive and management jobs is one which must be answered through empirical research. To this end, McCormick suggested this as a possible dissertation research project for me. This led to the development of the Professional and Managerial Position Questionnaire (PMPQ) in 1976. This is a 93 item structured questionnaire based in part on the PAQ but using 9-point rating scales (as is done in the Air Force program) and including several separate ratings scales.

This research has led to the identification of 10 job components common to many types of professional executives and managerial jobs. When tested via the prediction of salary levels, the part-of-the-job and complexity scales used in the PMPQ appear to be validated. Overall, a significant prediction of salary ($R = .68$, sign at p .05) was achieved using the ten components derived via principal component analysis. Currently, this instrument is being evaluated for refinement as a supplemental job analysis system but no final decision has been made. However, this research does have some implications for future studies of executive and management jobs.

One such implication is that it may be useful to obtain multiple ratings (part-of-the-job, complexity, etc.) of each job activity or area rather than rely on a single rating scale (such as importance or time spent). Such multiple ratings should be evaluated for a variety of purposes and specific regressions developed for each use.

A second implication is that there is probably a core of activities (tasks, duties, etc.) common to most management jobs. These probably include planning, budgeting, decision making, etc. The existence of such commonality suggests that job component scores may provide very useful data for definition, classification, and in making comparisons across job categories.

In defining training requirements and developing occupation-specific career development programs, the structured job analysis questionnaire may not be as appropriate as the Air Force task inventory approach which focuses analysis within a specialty or general occupation area. This implies that occupational analysts in the future may need to have preparation in more than one job analysis system and will have to be able to select the most appropriate technique to use for the particular purpose of a specific study.

This last implication is borne out by our recent experience in the Air Force occupational survey program. We have been tasked with several projects where our normal task inventory approach is clearly not appropriate. One such study was the Electronic Principles (EP) project which was reported in 1976 Psychology in the DOD Symposium. In this project an Electronic Principles Inventory of electronic fundamentals knowledges was constructed which permitted surveying across Air Force specialties.

Another example would be our present project to examine Air Force Professional Military Education (PME) as requested by the Commander, Air University. In this project we will need to build an inventory of professional knowledges in order to assess the Air Force requirement for enlisted and officer PME. This is a critical project since it has considerable curriculum implications not only for the PME schools but for all precommissioning programs as well.

The point here, is that a single job analysis system or methodology is not enough. A variety of techniques are needed to meet various requirements and in the Air Force, we have developed in this direction. We now have several job analysis appraoches and can select the most appropriate technique for a particular project. This to me seems to be one of the major significant trends in the Air Force occupational analysis program.

EVALUATION OF A HUMAN DEVELOPMENT PROGRAM FOR FIRST-TERM AIRMAN

Jack Miller
HQ Military Airlift Command

David Payne
Air Training Command

Paper presented at the Sixth Psychology in the DOD Symposium, USAF Academy, April 1978

In 1972 Military Airlift Command (MAC) initiated a project to identify the root causes of why first term airman (FTA) become involved in disciplinary actions. Preliminary investigations revealed a perceived problem concerning a lack of guidance, communications, and effective counseling between supervisors and FTA. Subsequent to these findings, an experimental program designed to provide additional counseling for newly arrived FTA and to provide squadron commanders a structured rehabilitation program for FTA receiving adverse actions was initiated at Dover AFB. Tentative findings of this experimental program indicated the program might reduce AWOLs, requests for separations, and improved duty performance (La Scala, 1976).

As a result of these findings, MAC instituted a command wide test of the HDP program. In conducting this test, four MAC bases were selected (Altus, McGuire, Norton, and Travis) to implement the program. The test bases were reponsible for conducting the test, while MAC HQs was responsible for evaluating the results. For comparative analysis, five non-test bases were selected: Charleston, Little Rock, McChord, Pope, and Scott. These bases became the comparison group for this study.

For purposes of this paper, only a very brief overview of the HDP test program will be presented. For further elaboration, the reader is referred to the final project report recently published by MAC HQs (1977). All newly assigned FTA were required to participate in a special two-hour orientation seminar conducted by trained counselors. The purpose of this session was to clarify and identify potential personal problems, and to refer the FTA to the appropriate assistance agencies if required (i.e., Finance, CBPO, etc.). The goal of this portion of the program was to identify and resolve personal problems of FTA when they first arrive on base and before they affect duty performance. The second part of the program involved only those airmen who had received their first adverse action as identified by their squadron commander. These airmen were entered into a structured program of group counseling and daily supervisor ratings. The goal of this portion of the program was to return non-productive airmen to full duty performance and to reduce subsequent actions which lead to administrative discharge.

The MAC HDP test was based on four assumptions: (1) the orientation seminar and questionnaire were effective in identifying FTA who were experiencing problems; (2) counseling these individuals would reduce subsequent adverse actions; (3) individuals entered into the structured rehabilitation program would demonstrate improved duty performance; and (4) those individuals in the rehabilitation program would experience a lower attrition rate than individuals not entered into the program. Based on these assumptions, we hypothesized that bases where the HDP was tested would experience an overall reduction in adverse actions in comparison to bases where the program was not tested. It was further hypothesized that these lower rates could be demonstrated between a select group of FTA exposed and not exposed to the orientation course.

METHOD

A one year (July 76 - June 77) time frame was established for the collection of base line data. No attempt was made to establish control groups at the individual bases, as all participating bases were required to employ the program in an "all or none" fashion. To reduce this group to manageable proportions, the following criteria were established. FTA were defined as those completing Basic Military Training or Technical School and arriving at their first permanent duty station; they must be assigned to a MAC unit on one of the test control bases; and they had to in-process during the time period 1 Aug 76 to 31 Oct 76. A total of 1386 FTA were identified. Of this number, 660 in-processed at the test bases, and 726 at the control bases. A record file was then established on these airmen. To evaluate the orientation program, airmen were tracked from arrival through 30 July 1977 (a minimum of 9 and a maximum of 12 months) at both test and control sites. The following criteria data were collected; Unfavorable Information File (UIF) entries, airmen separated for both cause and voluntary release, nonjudicial (Art. 15s) actions, and the number of court-marital charges preferred.

The second portion of the program (structured rehabilitation) was evaluated in a somewhat different manner. Project personnel at the last sites were instructed that all airmen who received a new UIF entry and who had less than four years service would be referred to the program. In addition to the criteria used to evaluate the orientation phase previously mentioned (Art. 15s, etc.), additional pre and posttest ratings on the FTAs measuring work performance were collected from their supervisors.

RESULTS

A summary of the data comparisons between test and non-test bases in presented in Table 1. All criterion data comparisons were tested individually at the 95 percent level of confidence. No significant differences were found among the comparisons, although differences in UIFs approached significance ($p=.10$).

Concerning results of the second portion of the program (structured rehabilitation), 79% of the FTA entered into the project were not involved in a subsequent adverse action. In an attempt to relate the cause of this reduction to the HDP, a comparison between test and non-test bases was made for adverse actions during the period of the test. This comparison, reported in rates per thousand, is presented in Table 2. No significant differences were revealed by this comparison. For these comparisons, criteria data were collected for all FTA's at test and control bases.

Supervisor ratings on FTAs in this program reflected a moderate pretest-posttest increase. However, the meaningfulness of this increase is difficult to interpret without a control group.

TABLE 1
COMPARISON OF HDP CRITERION DATA
TEST AND NON-TEST BASES

Percentages

Group	%UIFs	ART 15s	COURTS-MARTIALS	SEP FOR CAUSE	VOL RELEASE
Test (n = 660)	12.9	5.6	.5	3.9	2.3
Non-Test (n = 726)	15.2	6.3	—	4.1	1.4

TABLE 2
ADVERSE ACTIONS TEST CONTROL BASES
RATE PER THOUSAND
(1 Jul 76—30 Jun 77)

Group	ART 15s	AFR 39-10 (3-8L) Discharges	AFM 39-12 Discharges	AWOLS	Courts Martials
Test	32.9	16.8	9.9	4.4	1.9
Control	34.5	17.6	8.6	4.9	0.7

DISCUSSION

The results of this study failed to demonstrate a difference in the relevant criteria between test and non-test bases. Although the results of the 90 day structured rehabilitation program evaluation are encouraging, they must remain equivocal due to the lack of appropriate experimental controls and the fact that non-test bases reported the same low attrition rates without the structured program.

The question of why the MAC program did not produce the results predicted by the Dover experiment should be addressed. While we were able to duplicate the design of the Dover program and produce the the same uncontrolled results, we could not duplicate the time period of the Dover experiment (1973-1975). Between 1973 and 1977 numerous political, social, and economic changes took place in our nation that influenced the civilian community from which we draw our recruits. Also, the Air Force has changed. The slogans "People are Our Most Important Resource" and "Everyone is Our Business" were elevated from management cliches to a way of life. This emphasis on people programs produced numerous initiatives and modifications to existing efforts.

Another factor of importance when considering the results of this study is our data was collected for only a one year time-frame. This time-frame for the evaluation of a broad scale social program could be considered abrupt, as the effects might not demonstrable for several years (as reflected by "hard data" criteria). In effect the program may not have had time to prove itself. HQS MAC will continue to monitor adverse action rates command wide, and if a reversal in the current downward trend occurs a human development program for FTA may be reinstitued.

REFERENCES

LaScala, R. ***Behavioral services program final report. HQ MAC/ DPZ, Scott AFB, Illinois, 1976.***

Human development program test final report. **HQ MAC DCS/Personnel, Scott AFB, Illinois, 1977.**

FRAME THEORY AS A MODEL FOR TACTICAL DECISION MAKING

James R. Aldrich
Robert A Levit
The BDM Corporation

INTRODUCTION

There is a growing need to understand tactical decision making--the command, control and intelligence (C^3I) process--both as a basic determinant of overall force capability and as a specific functional component of a force. This need is inherent in the claim that C^3I is a "force multipler", i.e., a means to "leverage" the overall capabilities of a force to the point where it can achieve defensive objectives against a numerically superior enemy. Thus, the analysis and understanding of C^3I processes are important to force designers and developers in various respects.

First, the overall capabilities of a force are heavily dependent on its ability to effectively employ its resources in ways appropriate to the situation. To properly assess force design issues and development alternatives, C^3I ramifications must be addressed. For example, irrespective of a new sensor's performance characteristics, will the information it produces enable commanders to make decisions which are more timely or appropriate to the situation? Likewise, irrespective of a new weapon systems' capabilities, will it offer commanders new or better ways to pursue their battlefield objectives?

Second, the C^3I capabilities of a force are *themselves* subject to enhancement and degradation which will have a *leveraged* effect on overall force capabilities. Thus from a force design and development point of view, C^3I processes become candidates for improvement. Rapid advances in information processing technology--both hardware and software--offer powerful potentials in this area, but where do these potentials lead?

At present, it is difficult to answer such questions. Moreover, until C^3I processes are more rigorously understood in the context of total force capability, answers will be difficult to justify. Frame theory presents an approach to developing formal representations of C^3I processes and its application will hopefully contribute to a greater understanding of how these processes can be augmented to increase tactical advantage.

CONCEPTUAL BASIS: A PERSPECTIVE ON FORCE C^3I PROCESSES

The C^3I processes of a force encompass a series of continuous, closely related activities occurring simultaneously at all echelons of command which enable decision makers to consider essential information, assess the situation, make decisions, and communicate orders implementing those decisions. The system by which this overall process is carried out can be regarded as a network of individual C^3I elements. Each node (C^3I element) in this network represents the commander (and staff) or an organizational element of the force (Army group, air operations center, maneuver battalion, support command and so forth). Links in the network represent information flows among the C^3I elements including directions, status/intelligence reports, and requests.

The individual C^3I elements making up the overall network will vary widely in their specific information processing behavior due to their differing roles, responsibilities, and resources. Abstracting from specifics, however, many fundamental similarities appear. First, all C^3I elements have the same basic function: to efficiently utilize the resources at their disposal to accomplish missions assigned to them by higher C^3I elements. Individual roles, responsibilities, and resources specialize, but do not alter, this basic role. Second, all C^3I elements have the same basic problem: to perform their function on the basis of incomplete and possibly erroneous information. In other words, C^3I elements operate on the basis of their understanidng of the situation rather than the situation itself.

The notion of an *understanding of the situation (UOS)* provides a useful focal point for analyzing the information processing behavior of a C^3I element. A UOS can be regarded as a complex, dymanic structure of beliefs about the situation. In effect, it represents the "mental state" of a C^3I element and thus mediates between information inputs ("stimuli") and information outputs ("responses"): the UOS provides the context or "world view" against which new information is assessed, interpreted and, in some cases, translated into action. As such the UOS has a central role in the information processing of a C^3I element. Changes in the UOS provide "triggers" for decision making. Moreover, the UOS provides an information base upon which a C^3I element can actively attempt to "improve" its UOS (i.e., achieve a more complete or reliable correspondence with the "real" situation).

Focusing on the UOS and its role in C^3I processes, the information processing behavior of a C^3I element can be categorized into three basic types:

1. Processing concerned with developing, maintaining and updating the UOS as the "real" situation evolves through time.
2. Processing concerned with monitoring the UOS for operationally significant changes.
3. Processing concerned with structuring and resolving operational decision situations identified by the monitoring processes.

THEORETICAL FRAMEWORK: FRAMES AND FRAME SYSTEMS

The foregoing conceptualization provides a systematic decomposition of force C^3I processes into a network of specialized, interacting C^3I processes, and proposes a structure of subprocesses concerned with using a UOS to perform the basic C^3 function. The conceptualization will be useful, however, only to the extent that it can be translated into a more rigorous framework permitting critical analyses and experimentation. There is a growing body of work which can be brought to bear on this translation. This work has developed in the "cognitive sciences" at the borderline of psychology, linguistics, and artificial intelligence, and may be broadly characterized as building "information processing theories of mental behavior". Heavily oriented towards simulation-based representations, this work is especially pertinent to the problem of introducing C^3I processes into conflict simulations.

One of the conjectures which has emerged from work in the cognitive sciences is that much of human knowledge is organized and utilized in terms of hierarchical systems of information and procedures for recognizing and responding to stereotyped but relatively particularized situations. This view interprets "understanding" as finding and fitting the "right" stereotyped information and procedures to the situation (rather than assessing and applying a succession of broad "general principles"). Whether or not this conjecture reflects the actual organization of knowledge in the human mind, it has proven to be a useful working hypothesis in developing simulations which exhibit, in a limited sense, "understanding".

Among the various simulation approaches which have emerged along with this conjecture, one of the most direct is based on the concept of a *frame*. To quote from a seminal paper on the frame concept: "Here is the essence of the theory (of frames); When one encounters a new situation. . .one selects from memory a substantial structure called a frame. This is a remembered framework to be adapted to fit reality by changing details as necessary."[1] In essence, the, a frame is a complex information structure representing a stereotyped situation or activity, *together with* procedures for "matching" or "fitting" the stereotype to a range of particular situations.

Originally developed in the context of visual scene analysis, the frame concept has since been extended to broaer types of situations.[2] It can be extended to the context of C³I processes in the form of a *decision frame*, i.e, a special type of frame which represents a stereotyped decision situation and outlines the procedures by which a C³I element deals with it. Decision frames may be further subdivided into two basic types: (a) *situation frames*, representing decisions about the nature of the tactical situation ("What's happening?") and, (b) *action frames*, representing decisions about how best to respond to the tactical decision ("What should I do about it?").

A C³I element's situation frames provide the means by which it interprets the information it receives as input. Each situation frame describes a stereotyped situation and is intended to represent that C³I element's approach to recognizing the occurrence—or impended occurrence—of that situation. Thus, for example "penetration" and "high nuclear threat" might be typical situations described by situation frames.

A situation frame will generally include such procedural information as: (a) conditions under which the frame is to be activated; (b) a description of the situation in terms of key features and their characteristic relationships; (c) procedures to control the application of the recognition procedures, integrate their results, and take further action as appropriate.

Once activated, the recognition procedures of a situation frame are applied to the information received by the C³I element. This process is controlled by the assessment procedures which then receive the results of the applications and decide what action to take next. Thus, it is the assessment procedures which ultimately determine whether or not the situation is "recognized" and, if so, what should be done about it.

A C³I element's action frames represent its "perspective on the situation" in terms of its objective, the decisions it has made and their expected effects, and the types of decisions it may have to make to keep the situation evolving in the direction of its objectives. Each action frame "outlines" a particular decision process and is intended to reflect a C³I element's approach to dealing with a generic kind of decision problem. Thus, for example, a C³I element may contain "attack", "defend", and "delay" action frames.

An action frame contains such procedural information as: (1) activation conditions under which the frame is to be operative; (2) types of situation-related information needed to conduct the decision process ("essential elements of information" associated with the decison); (3) procedures for developing relevant alternative courses of action; (4) procedures for evaluating alternative courses of action considering the objectives and possible enemy reactions; (5) criteria for selecting among alternative courses of action considering the evaluations; and, (6) procedures for implementing the selected course of action.

The decision process represented by a particular frame is initiated when its activation conditions are satisfied. Currently available information is screened against the frame's essential elements of information; relevant information available is "read into" the frame while requests may be generated for unavailable items. Alternative courses of action are developed and evaluated in terms of their influence on enemy reactions. The "most promising" alternative is selected and translated into an appropriate operations order. The frame remains in operation until the objective is met or until the initiating conditions cease to hold (e.g., the mission is changed).

The power of frames derives from their ability to activate other frames. This permits a hierarchical decomposition of "complex" decisions into "simpler" decisions. It also permits very complex recognition-response behavior to be represented within a system of frames. As an example, the activation of a broad defense action frame might well result in the activation of a penetration situation frame for the purpose of "watching for" impending or developing penetrations. However, once a penetration was recognized by this situation frame, it might in turn activate a penetration response action frame to determine what to do about the penetration. Notice that the activation of the penetration recognition frame reflects the identification of penetration as a significant contingency.

Frame theory itself is formally neutral with respect to the contents of the frames and their interlinkages. Thus, for example, situation frames may involve very simple yes-no conditions or very complex pattern recognition mechanisms; likewise, action frames may be based on simple heuristics or complicated decision-theoretic procedures. The virtue of this neutrality is the greater flexibility of frames as a vehicle for modeling and analyzing C³I processes. Indeed, using frames in the C³I context will involve developing particular systems of situation and action frames, assessing their behavior (against, e.g., a man-in-the-loop's behavior), and modifying them accordingly. The price of the neutrality is the lack of theoretical guidance as to what frames should be included in the system, how they should be formulated, and how they should be linked.

IMPLEMENTATION GUIDELINES: THE DOCTRINAL COMMANDER

This guidance can be supplied by doctrine. Quoting from Army Field Manual 100-15, *Larger Unit Operations*, ". . . doctrine is defined as enlightened, tentative statements of what will usually work best. . . . Doctrine should indicate and guide, but not bind in practice. Its application requires judgment in adapting to the peculiarities of the situation. . .". The correspondence between this characterization of doctrine and the earlier characterization of frames is readily apparent: there is a close analogy between the role of doctrine in the "real world" and the role of frames in the theoretical framework. Accordingly, a reasonable implementation strategy is to begin by developing a doctrinally-based system of frames representing, in essence, a C³I element which operates in strict compliance with doctrine, i.e., *a doctrinal commander*.

Methodologically, the postulation of a "doctrinal commander" is analogous to the postulation of the "rational man" in economics or the "idealized speaker" in linguistics. In all cases, the aim is to provide some external standard which can be assessed *independently* of individual differences. The doctrinal commander is unlike "rational man" and "idealized speaker" in several respects, however. He is

[1]Marvin Minsky, "A Framework for Representing Knowledge," Chapter 6 in P.H. Winston (ed), *The Psychology of Computer Vision*, McGraw-Hill, 1975, p. 212.

[2]See work reported in D. Babron and A. Collins (eds), *Representation and Understanding: Studies in Cognitive Science*, Academic Press, 1976, and R.C. Schank and B.L. Nash-Webber (eds), *Theoretical Issues in Natural Language Processing*, Bolt, Beranek and Newman, 1975.

not necessarily "best" or "ideal". Moreover, he is not nearly so well defined: (a) doctrine does not provide detailed rules for every conceivable situation; (b) there may well be inconsistencies among different doctrinal sources (training literature, field manuals, etc.); and (c) doctrine is constantly changing.

In spite of these difficulties, the doctrinal commander appears to provide the most explicit external standard currently available in the C^2I area. A well developed "doctrinal commander" also offers a natural point of departure for considering individual differences in C^2I processes.

POTENTIAL APPLICATIONS

The approach presented in this paper has evolved in the context of the design and development of comprehensive conflict simulations. We are currently designing a series of "doctrinal commanders" along the lines indicated above to represent C^2I processes in a conflict simulation focusing on decision making at echelons above division. In addition to their modeling and simulation applications, however, these concepts may be useful in the design and development of automated information systems to assist commanders (so called decision aids). For example, there are marked similarities between the frame concept and the tactical templating concept as described by U.S. Army Field Manual 100-5, *Operations*. Frames may therefore provide a powerful means to assist intelligence officers in formulating and using tactical templates to assess tactical situations.

THE NATURE OF THE NAVIGATOR AND RADAR NAVIGATOR CREW DUTIES ON A B-52 MISSION

Earl D. Sharp and Gerald P. Chubb

INTRODUCTION

The B-52 was originally designed in the early fifties as a high-altitude bomber. The G model was manufactured in the early sixties. By the mid-sixties, high altitude surfact-to-air missiles made it necessary to consider low-level flight as a viable strategy for surviving the penetration portion of a mission.

The B-52G has a six man crew consisting of three two man teams (T.O. 1B-52G-1, 1975). The pilot and copilot are responsible for aircraft maneuvering and altitude control, engine and fuel management, supervision and control of hydraulic/electrical power, and environmental control. The defensive team consists of an electronic warfare officer and an enlisted gunner seated side by side, facing aft. The offensive team consists of a navigator and radar navigator seated side by side facing forward.

The primary job of the offensive avionics team is to stay on-course/on-time and deliver the weapons payload against their pre-assigned targets. Each bomber must stay within prescribed limits of the flight plan in order to: (1) assist other attacking bombers (in suppressing threats), and (2) avoid the weapons effects from other planned attacks.

The radar navigator is the senior member of the bomb/nav team and is principally responsible for radar fixes and keeping the bombing navigation system (BNS) updated. Errors induced by sensor measurement bias and noise as well as the accumulated integration error in the analog computers must be periodically corrected by fix taking. The radar navigator is also the lead operator in all gravity weapons releases.

The navigator is the lead operator in all missile launches, (e.g., the Short Range Attack Missile—SRAM). The navigator's principal responsibility is heading management: seeing that the best available heading information is provided to the BNS computer. There are four different sources of heading information: (1) magnetic heading (N-1), (2) true heading (AN/AJA-1), (3) MD-1 Automatic Astrocompass true heading, and (4) SRAM inertial true heading. All but SRAM inertial heading data can be fed directly into the BNS.

RADAR NAVIGATION

AFM 51-40(1973) describes all of the basics for navigation. There are five repetitive steps involved in the navigational process (CDC 34350, 1969): (1) determining where you are at the start of the navigation leg, (2) determining the next destination point, (3) determining how fast you are going and in what direction, (4) keeping a constant account of where you are now, and (5) determining the corrections necessary to stay on-course and on-time. The BNS should be an aid in each of these steps, but should not be relied upon since this results in the risk of getting lost when the hardware malfunctions. A sense of mental reckoning must be developed as a cognitive skill. Radar scope interpretation (RSI) is a key element in acquiring this skill. An experienced operator can use his radar display to validate the coarse accuracy of this instruments, even though they may provide more precise information when working properly. There is no other single skill of greater value nor any sensor of more importance for navigation.

At high altitude, a large area can be viewed, but depending on the radar range selected, there are scale differences between the radar display and the navigator's map. Judgment and evaluation are required in interpreting the range(s) and bearing(s) to (and among) returns on the display versus radar-significant features on the map.

Radar tuning that makes towns and other cultural returns show up better typically washes out terrain features, and conversely. The tilt of the antenna will affect the path of the radiated energy and, in turn, the quality of the imagery displayed. This and the receiver gain control, which can affect how much spurious clutter (or noise) is displayed along with valid information, are only controllable at the radar navigator's station. The radar navigator also has exclusive control over the range scale, markers, and format presented on the display.

PENETRATION

After receipt of a valid "Go-Code", the nav team will execute their Weapons Preparation checklist. Prelaunch and Strike checklists are required before missile launches, and Pre-IP and Configuration checklists are required prior to gravity weapons releases.

Descent to low level requires completion of the before and during descent checklists by both pilots and offensive systems operators. The radar can be used to perform a high altitude calibration, prior to descent. Malfunctions or other distractions during descent can draw attention away from the altimeter. As lower altitudes are reached, certain equipment needs to be set up to new operating modes. (Personal locator beacons are unsnapped, zero delay lanyards are fastened, the radar altimeter becomes functional,and the radar sets needs to be set up for operation in a terrain avoidance mode). From this point on, the radar navigator cannot use tilt to improve his imagery without affecting the pilot's terrain trace display.

Radar range at low level is restricted. Terrain features that preclude detection by enemy tracking radars, also create shadows on the radar display that can hide significant ground imagery used for RSI. Turns at low altitudes also affect what will be displayed on the radarscopes. The offensive team's displays provide useful information at greater distances and wider fields of veiw than the pilot's displays. The offensive team, therefore, becomes responsible for alerting the pilots to terrain features they need to be looking for on their displays. It is also important to alert the pilots of the range and bearing to known man-made obstacles and to pre-brief the pilots on the terrain elevation along the anticipated flight path. This is especially important at low altitude and bcomes a dominant part of the navigation team's secondary task loading. Vibration at low level is a significant factor and aggravates reading displays (Speakman and Rose, 1971).

At low level, it is also necessary for the navigator to update the Fuselage Reference Line (FRL) entered into the BNS. As fuel is used, the aircraft gross weight changes, and the angle of attack will change. This alters the attitude of the radar sensor platform as well, so the navigator's FRL entry is effectively a compensation that must be kept updated if the radar navigator's display and pilot's terrain trace are to provide valid information.

Threat encounters that seriously endanger the survival of the aircraft obviously jeopardize the mission, but any response to maneuver the aircraft out of imminent danger has an impact on navigator workload and, in turn, also affects mission success. The aircraft will be driven off-course and will fall behing schedule. These maneuvers, along with any compensating attempts to catch up with the flight schedule deplete fuel reserves needed for post-strike recovery. These changes in fuel usage impact FRL updates.

STRIKE

The bomb run itself is a period of peak workload for several reasons. First, the operator's attention level is raised by the impending weapons release, but he cannot affort to miss any indication of equipment malfunction that might signal the need to adopt some alternate bombing technique. Second, he must accomplish all those tasks that set up these alternates, if they need to be employed, without compromising his pacing for the planned approach for the present target. Third, target areas may be heavily defended, and he must decide when to request a cessation in any on-going evasive maneuvers if these preclude his seeing the OAP. Fourth, maximum target destruction and minimal collateral damage rest on his being able to aim precisely.

The problems associated with aiming are formidable. First, the radar navigator must depend on the navigator having done his job: they know where they are, where they are going, and have compensated for all known systematic errors in the BNS. Second, the radar navigator must be able to pick out or acquire the offset aiming point. Cultural features at one range may take on a form that is appreciable different from their appearance as range diminishes. As some of the more distinct features break-up, disintegrate, and change form, other nebulous features become sharper, more refined and brighter as range decreases. Recall that radar tuning also affects what kind of features stand out on the display. Also note that spurious returns will come and go. Scintillation, glint, and cardinal point effects all create irrelevant information. Enemy jamming can further clutter the display, reduce contrast and generally harras an already heavily burdened operator. Third, the radar navigator must not only do his own job, he must be sure the navigator stays out of trouble and that the pilot doesn't compromise the bomb run by trying to maneuver at an inappropriate time or try to follow a Flight Control Indicator that is being fed bad data. Finally, having acquired the offset, the radar navigator must move the crosshairs onto the point only if he is sure the displacement is due to BNS error and not some oversight of his own. Otherwise, attempts to refine the crosshairs may in fact be driving the system away from the target. Errors of this sort are dependent on the geometry of the situation (target and OAP locations), the procedures employed by the operator, and the workload stress prevailing at the time.

In the case of a damage assessment sortie, all of this must be accomplishd in the face of uncertainty. Until the aircraft gets close enough to determine whether or not the damage inflicted by a previous strike was adequate, the radar navigator must delay his decision to expend or withhold the weapon. By the time he is close enough to see the target area well enough to evaluate the damage, the decision must be made quickly. To allow time for handling all the contingencies that might arise with equipment malfunction during this critical period, the operator must be well trained and self-disciplined.

Weapons releases do not always occur as single, isolated events, and there are occasions where missile launches can occur in proximity to gravity weapons releases. Because of these missions requirements, the workload of the operators may be positively correlated. When a gravity weapon is released first, full attention can still be dedicated to the missile launch, since malfunctions can force downgrading of the weapons release but should not compromise the bomb release point. However, the missile has a launch window and can go into an extended hold. If problems are encountered in diagnosing and ~~o~~r correcting these difficulties, the time used must necessarily detract from the time available for the subsequent weapon release. Consequently, a missile launch followed by a weapons release can present a more complex task than a weapons release followed by a missile launch.

POST-STRIKE RECOVERY

Once all weapons are expended, the final post-release checklists are accomplished and low-level navigation continues until the aircraft clears the enemy threat area. The aircraft will then climb to a fuel-efficient cruise altitude. Recovery to a friendly post-strike base can entail a variety of contingency decisions depending on fuel reserves, condition of the aircraft, proximity of various resources for regeneration (for subsequent sorties), and a variety of command ~~c~~ontrol factors. The nav team may have to engage in some degree of enroute mission planning to successfully recover from the strike mission.

CONCLUSION

Despite all of these difficulties, the experienced crew has an ability to adapt to circumstances in a manner not achievable by even the best automated system. When the equipment fails entirely, the crew can still accomplish a weapons release by alternate means, and a wide variety of techniques apply between the extremes of fully operational and completely inoperative equipment.

REFERENCES

AFM 51-40, *Air Navigation*, Government Printing Office, Washington, D.C., 1 Jul 1973.

CDC 34350, *Navigation /Bombing Trainer and Flight Simulator Tactics Specialist*, Vol 3, Extension Course Institute, Air University, Gunter Air Force Base, AL, 1 Apr 1969.

Holloway, Bruce K., "Thoughts on TRIAD," *Aviation Week and Space Technology*, Vol 98, No. 3, 15 Jan 1973, p 9.

Speakman, Jerry D. and Rose, Justus F., *Crew Compartment Vibration Environment in the B-52 Aircraft During Low-Altitude, High-Speed Flight*, AMRL-TR-71-12, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, Mar 1971.

T.O. 1B-52E-1-6-1, *Radar Navigator's /Navigator's Manual, Inflight Operation, AN /ASQ-38 Weapons Control Systems-Offensive*, 1 Aug 1964.

T.O. 1B-52G-1, *B-52G Flight Manual*, 1 Jan 1975.

HUMAN ENGINEERING PROBLEMS IN THE B-52 OFFENSIVE AVIONICS CREWSTATION

Gerald P. Chubb

INTRODUCTION

In any system development effort, there are time and cost constraints that limit the scope and depth of human engineering applications during design and development. Typically, the front station design (pilot/copilot) receives primary attention. Other crewstations receive attention as time and dollars permit.

Once an Air Force system has been acquired, management responsibility is assumed by Logistics Command (AFLC) and assigned to an Air Logistics Center (ALC) where system management (SM) and item management (IM) responsibility is exercised. Currently, AFLC has no organic human factors engineering capability. The SM has the authority to make modifications to the system, so long as these do not entail significant amounts of research and development. New panels added to the crewstation will be placed where space is available and with a goal of minimizing the attendant costs.

The Strategic Crewstation Design Evaluation Facility (SACDEF) was established when the electronic warfare community requested an assessment of crew workload prior to the Phase VI modifications to the B-52. Based on the success of this effort (87 of 93 human engineering recommendations were immediately incorporated), SACDEF was expanded to provide comparable support to the navigation community.

AFLC has utilized SACDEF study results to accomplish the Phase VI modifications to the electronic warfare officer's crewstation on the B-52. They have been a major proponent for similar SACDEF studies in support of the proposed Offensive Avionics System update required to meet Strategic Air Command's operational requirements. Oklahoma City ALC and the contractor both were given the raw data and the statistical analysis of the questionnaire survey results upon which this paper is based.

METHOD

In preparation for a Baseline Study of navigator and radar navigator workload in the B-52G, a Preliminary Study was conducted in February 1977.

Subject Selection

A questionnaire survey of the combat ready crew force was conducted. A systematic (non-random) sampling procedure was used (Hoffman and Chubb, 1978) to obtain a representative sample of six crews. Sample size was limited because this was a pre-test of materials and procedures to be used in the Baseline Study (involving 24 crews).

Procedure

Crews were sent TDY to SACDEF to participate in the study. The first half-day was spent in an orientation briefing. The second half of the day was spent studying a Combat Mission Folder (CMF) comparable to a Single Integrated Operating Plan (STOP) Emergency War Order (EWO) sortie that would be reviewed by a crew on alert. The following morning, the crew executed a real-time simulation of this mission in a modified AN/ALQ T-10 Bombing Navigation Training Simulator. Instructor personnel were assigned TDY to observe/critique the nav-team performance and to play the roles of pilot and copilot over interphone. The next day was spent in the trainer conducting an open-ended, structured interview. These sessions were taped and subsequently transcribed. The solicited crews remarks then served as the basis for the construction of an rating-scale questionnaire that was administered during the Baseline Study. Since these responses are still being analyzed, the present paper is based on the Preliminary Study results.

RESULTS

The single, largest driver of workload is unreliable equipment. Crews can expect to have to "work around" one or more malfunctions every time they fly. There is no single, integrated malfunction indicator or caution annunciator panel as would typically be found in the pilot's crewstation. Those malfunction indications that do exist are dominantly placed in peripheral viewing areas where neither operator can see them. Airspeed and altitude deviation indicators for gravity weapons release are also scattered. While the Short Range ttack Missile (SRAM) deviation indications are relatively centralized, the navigator's C-4 table lamp is usually positioned where it obscures these indicators. Raising it out of the way would then obscure the navigator's radar scope.

The other major defect in the basic concept of the equipment design is the inherent need for the operator to accomplish one set of steps for training missions and another set for combat missions. Inhibition effects may induce a negative transfer of training effect. Crews acknowledge an awareness of this problem. For example, some expect to attack an enemy target and inadvertently "turn the tone on" as they would on an Operational Readiness Inspection (ORI) mission. In peacetime training this allows them to be evaluated by a Radar Bomb Scoring (RBS) site to determine estimated CEP, but in wartime missions this tone could reveal the location of an otherwise undetected aircraft.

Radar Navigator Station Comments

The ballistics derivatives for gravity weapons cannot be set by the radar navigator (N) nor can he set up his radarscope recording camera. Both must be set by the navigator (N). While the RN can set time of fall and trail, the indicator panels are too far away to be read accurately especially during low level flight. None of the navigation data are readily accessible, and major changes in body posture must be made (against the inertial reel of the restraint harness) to read these displays. Even then parallax is a problem. The radar pressurization panels are on the navigator's right console and may be obscured from the RN's view depending on the adjusted seat positions and the N's body position. Malfunctions can go undetected until the radarscope presentation reflects significant performance degradation.

The radar failure warning light is behind the RN's left shoulder, and should the radar receiver gain malfunction, the only control the RN has over his radar display gain is via the AJ-1 gain control located on this same panel. Taller personnel reach under the armpit, but shorter operators have to reach over their shoulder to make these adjustments. Cramping quickly results, in the latter case especially. It is not uncommon (especially in the G model aircraft) to need to recenter the radarscope display after turns, and these controls are inconveniently located next to the terrain test panel where the AJ-1 gain control resides.

The interphone panel in the G-model aircraft is inferior to the H model because only one volume control is available. When weak signals are made audible, strong signals become deafeningly loud. The panel is somewhat inconveniently located, especially for adjustments needed to communicate outside the aircraft (for refueling, landing, weather reports, etc.).

Navigator Station Comments

The most serious problem is the location of the system power switch for SRAM. Experienced crews, intending to reset the AUTO MANUAL switch, have instead shut down power, sometimes just prior to a practice launch. Field modifications have been made to discourage inappropriate operation.

Navigator's are jealous of the natural, "built-in" storage space the RN has over his topocomp display. He can put his checklist and CFP materials up there. The navigator has relatively more paperwork and less space for storage. If the oxygen bottle is mounted in a forward leaning direction, he cannot even stow his nav-bag to the right of his seat without covering a heating duct.

The navigator has no control over receiver gain, range scale, nav/bomb markers or radar presentation format (PPI vs. OC). He must request (and often wait for) the display he needs for navigation. There is no way for him to use the radar crosshairs nor control them. His radarscope display is only half the size of the RN's 10-inch topocomp.

The heading select switch is unguarded and positioned low enough that it can be inadvertently bumped and set to the wrong position. The doppler-indicator lights are low enough that a box lunch will preclude their being seen by the RN, but the doppler switches are up high where the shorter navigator may have trouble reaching them. The APN-69 malfunction indicator and controls are also less than readily accessible.

While the navigator is responsible for timing on an alternate bomb run, the time-to-go indicator is at the RN station. These indicators require the RN take his head (eyes) "out of the scope." There is a parallax problem in reading heading error, even for the RN much less the N. The Electro-optical Viewing System (EVS) indications of these FCI values are so inaccurate they remain unused. Only the radar altimeter display is valued. The Indicated Airspeed Display would also be used if it were more accurate.

Other Equipment Comments

The floodlights are not used because of the reflections induced on the topocomp. Some circuit breakers are virtually unreachable unless the operator unstraps. The thumbwheels incorporated in the Automated Offset Unit (AOU) design are just backward from those incorporated in SRAM. Consequently the operators find themselves discovering by trial and error which digit the thumbwheel actually controls. The card reader used for entering AOU data is awkward to use. However, the AOU itself has been well received and is being used in a mode not anticipated during design that tends to facilitate accurate navigation: setting turn points up as "dummy" targets. This device has sufficiently unburdened the N that he can now "kill" heading error (on direction from the RN).

Other Complaints, Suggestions, and Recommendations

Mission folder materials may include copies of maps, and crews note that terrain features are often unreadable. The folders are laid out in a manner that is economical, complies with directives, and is relatively easy to prepare, but unfortunately, the result is a package that is awkward, bulky, and difficult to use. Current conventions in target labeling result in a double numbering scheme that can lead to confusion and operator error.

Digital calculators are used by a few crews, and quite effectively. However, many crews suggested they prefer their MB-4 circular slide rule: (1) it has no batteries, (2) it will not break if dropped, and (3) they already know how to use it.

Many of the crews had participated in Red Flag exercises at Nellis Air Force Base. Still, they found the SACDEF studies were beneficial from the crew's standpoint because the realism of the SIOP/EWO sortie simulation made them aware of a number of aspects of the combat mission that cannot be fully duplicated in Red Flag. No single training experience fully prepares a crew for combat, but a variety of exercises can expose crews to the different considerations they need to be aware of if they are to be prepared for executing the SIOP.

Many of the more experienced crews exhibit a tolerance of the status quo because they have adapted to their work environment. Less experienced crews are more critical. Recommendations made by the less experienced crew member are often not supported by the more experienced crew member. Crewstation design and evaluation based on the opinions of only experienced crew members could very easily aggravate proficient performance from less experienced crews. The current SAC crews are relatively less experienced than those of five years ago (much less those of a decade ago).

CONCLUSIONS

The Offensive Avionics System update is bound to reduce crew workload just by digitizing the system and improving equipment reliability. Additional gains can be achieved by rearranging retained equipment panels. User acceptance of proposed changes can be facilitated by conducting SACDEF studies. Not only do these allow empirical and quantitative assessments of performance improvements, they allow substantive crew comments to be made. The impact on crewstation design is twofold: crewmembers can influence the equipment they (and/or their successors) operate, and they can be told the rational bases for the design trades that led to the proposed design.

REFERENCE

Hoffman, Mark S. and Chubb, Gerald P., *Sampling Methodology Developed for Preliminary SACDEF Study*, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio (in preparation).

PERCEPTUAL AND MOTOR ORGANIZATION OF SEQUENTIAL PATTERNS

Richard E. Christ
New Mexico State University

An examination of previous research on perceptual-motor performance reveals an interesting but unfortunate gap in our understanding of this type of information-processing skill. Limiting this discussion to the situation in which a subject is required to respond to a temporal pattern of stimulus inputs, it appears that at least two different types of analysis have dominated the experimental literature. Furthermore, it appears that there has been little attempt to integrate these two approaches.

On the one hand, there has been considerable effort exerted to develop an understanding of the variables which influence the perception of temporal patterns. The bulk of this research has dealt with visual patterns, but some recent work has been concerned with the perception of temporal patterns presented in several different sensory modalities (Handel & Buffardi, 1969). The results of these studies generally support a Gestalt interpretation of temporal pattern perception; *higher-order invariants of individual elements were the important perceptual* variable (cf., Gibson, 1969). More specifically, these studies demonstrate that an analysis of pattern structure could lead to accurate prediction of the subject's organization of the patterns and the relative difficulty in perceiving the patterns. Furthermore, while easy patterns tend to be easy and identically organized for all sensory modalities, difficult patterns were not. Handel and Buffardi (1969) showed that the performance of pairs of modalities depended upon the inherent properties of the individual modalities.

These experiments on perception have not systematically analyzed the effects of response factors on perception. That response factors have an effect has been documented. Royer and Garner (1966) pointed out in the discussion of their article that ". . . after the perceptual organization has occurred, the pattern could be responded to in complete synchrony with little difficulty . . . . Forced responding too early, particularly with the more difficult sequences, leads to almost complete disorganization, and actually interferes with the perceptual organization process" (pp. 45-46). Royer, in a personal communication, also pointed out that "a number of subjects . . . displayed a variety of body motions such as head-bobbing and bodyrocking during the period in which the perceptual organization has been established." Royer felt these motions could have been used to facilitate the organizing process or to help retain an already perceived organization. A recent review of the literature on human information processing and sensory modality effects concluded that with spatial or temporal patterns, there is an interaction between modality and information complexity and that generalizations about modality-information interactions depend on the control of response factors (Freides, 1974).

On the other side of the perceptual-motor performance issue, those investigators who have examined motor skills or motor control have largely ignored the subject's perception of the pattern to which he was responding. For example, in tracking tasks, the characteristics of the input function and the method of displaying this input have been varied but there has been little attempt to determine what the subject perceives or how tracking performance might vary as a function of the subject's perception. Several recently conceived models of voluntary perceptual-motor performance have incorporated pattern perception and schema generating features (cf., Pew, 1974). McGee and Christ (1971) showed that a redundant auditory input could either facilitate or inhibit performance in an otherwise visual tracking task, depending upon which specific aspects of performance was being monitored. Even though these and other evidences exist to underscore the importance of pattern perception, there has not been a concerted effort to examine these effects in experimental studies of motor behavior.

Experiment I

In experiment 1, the subject performed a pursuit tracking task. The individual elements of the step-function input appeared as a straight vertical line which moved in discrete jumps from one position to another along the x-axis of the 10 cm face of an oscilloscope. By moving a lateral arm control, pivoted at the elbow, a subject controlled the position of a similar vertical line. His task was to keep the two lines superimposed.

The input function was a pattern composed of eight positions which was replicated without a break eight times per trial. The target remained in each position for .8 sec before jumping to the next position. Any given input function could be identified by a sequence of eight digits which corresponded to the eight positions number 1 through 8 from the extreme left position to the extreme right position. Each sequence of positions could be organized into eight different patterns, corresponding to the eight possible starting points in the sequence. This study used two different organizations of one pattern: Sequence A in which the sequence of positions was 4-3-7-1-8-5-2-6, and Sequence B in which the sequence of positions was 8-5-2-6-4-3-7-1. The latter is merely Sequence A reorganized so that it begins with the fifth element in Sequence A.

Six groups of subjects were employed in the experiment. Two control groups were presented with Sequence A or Sequence B for 25 trials on each of two days. Two transfer groups (P-P) were trained for 35 trials on either Sequence A or Sequence B and then, at the beginning of Trial 11 on the second day, were transferred to the other sequence. Two additional groups (D-P) of subjects were given a stimulus input which consisted of the same organization of binary left-right direction of movement as were found in Sequence A and Sequence B, respectively, but with the actual target positions chosen at random from those available in the specified direction. For example, the pattern of left-right directions in Sequence A (L-R-L-R-L-L-R-L) could be generated by the position sequence 3-2-8-5-7-4-1-6 followed by the position sequence 2-1-5-3-8-6-4-7. On Trial 11 of the second day, these subjects were transferred to the fixed pattern of target positions (Sequence A or B) having a new starting point in the pattern of left-right directions.

Since data analysis showed no effect due to sequences, the data were pooled to show only the effects due to the control, transfer of positional sequences (P-P), and transfer from directional to positional sequence (D-P). The results show a relatively stable level of performance maintained by the control subjects (about 14 volts of integrated error) but a large (significant) increase in tracking error for both (from 13 volts to 15 volts for P-P and from 21 to 23 volts for D-P). The performance of subjects in the transfer conditions improved rapidly over the trials immediately following transfer (from 23 to 15 volts for P-D). An analysis of oscillographic records of target and cursor positions showed that the larger global measures of tracking error were associated with errors in response timing. On trial 10 the P-P and control subjects were correctly leading (or anticipating) the target about 90 percent of the time while subjects with the directional pattern were leading the target only about 15 percent of the time. When the starting point (organization) of the input was changed, the P-P subjects show a marked decrease in the percentage and magnitude of anticipatory leads. Both the P-P and P-D subjects show a predominance of anticipatory movement by the end of their exposure to the new organization of position inputs.

To assess the subject's perception of the input patterns they were asked at the end of both days of tracking to describe with a paper-and-pencil test their organization of the pattern they were tracking. The format of the test consisted of eight consecutive frames with a row of eight vertical lines in each frame. Subjects were asked to circle the position assumed by the target over successive steps in the input pattern. Quickly summarizing these data, no more than one-third of the subjects in the fixed positions input groups could correctly describe the patterns they were tracking—only three of the directional pattern subjects could correctly report the pattern of directions. After transfer, only 5 out of 16 subjects could describe the pattern organization to which they were transferred. The fact that *all* subjects show by their anticipatory movements that they can produce the pattern of movements required but that less than one-third can describe the pattern suggests that the subjects' perception of the pattern and their responses to it were independent. Furthermore, a significant correlation between pattern identification scores and overall tracking performance suggests that if subjects perceptually organize the pattern the same way the experimenter organizes it, they performed better than if their perceptual organization differed from the objective organization.

Experiment 2

Experiment 2 was designed to more specifically examine the perception of temporal patterns, patterns like those used in tracking tasks. The temporal pattern in this case was a six-position pattern which was replicated without a break while the subject passively viewed its presentation. As before, the stimulus remained in each position for .8 sec before moving to the next position in the sequence. When the subject thought he could describe the pattern he pushed a button which stopped the presentation and he then described it in one of three ways to be described below. The stimulus patterns consisted of roughly one-half the 720 patterns which are possible in a sequence of six events. If the positions are numbered 1 through 6 from left-most position through the right-most position, one very easy pattern could consist of the sequence 1-2-3-4-5-6-1-2. . . This pattern may also be represented by five short movements to the right followed by one longer movement to the left. Equivalences among the 720 different sequences can be defined by varying the starting point of the sequence of position or direction components and by generating complementary patterns. Hence, the sample pattern given above generates 11 equivalent patterns when starting points are varied (e.g., 2-3-4-5-6-1, 3-4-5-6-1-2, etc.) and when the complement (or mirror images) of the patterns are generated (e.g. 6-5-4-3-2-1 is the complement of 1-2-3-4-5-6, etc.). There are 56 basic patterns like the example given above, each of which generates 12 equivalences by considering starting points and complements. Another 8 basic patterns generate only 6 equivalences each since the complement of these patterns, began at one starting point, was a pattern already defined by another starting point (e.g., the complement of the pattern 1-2-3-6-5-4- is 6-5-4-1-2-3- which is the original pattern begun at the fourth position).

Each of the 84 subjects viewed all 64 of the basic patterns, 32 patterns on each of two days. Six subgroups of 14 subjects each began each pattern at a different starting point. Thirty subjects were required to describe the patterns by naming the temporal sequence of position; for these subjects the six positions were numbered consecutively from left to right. They described the patterns by naming the position which occurred first in the pattern, then the position which occurred second, etc. Another group of 30 subjects used the paper-and-pencil test described for Experiment 1. The format of the test consisted of six consecutive frames with a row of six vertical lines in each frame. The remaining 24 subjects were asked to describe the pattern in terms of the sequence of left-right movements of the target over successive steps in the pattern; they were specifically told to disregard the position of the target between movements. The objectively defined starting point for each pattern and the order of presenting the 64 basic patterns were balanced over subjects.

Each subject's performance was examined as a function of three dependent measures. Pattern difficulty was presumed to be reflected in the number of target elements presented before the subject terminated the presentation and in the number of subjects who made an error in describing the pattern. The subjectively defined starting point of patterns that were correctly described was used as a measure of pattern organization.

At the time of this writing, the data from Experiment 2 were still in the process of being analyzed. However, some very preliminary results are worth mentioning. It has been possible to rank order the 64 basic patterns by their relative difficulty. Hence, the pattern 1-5-2-6-3-4 was correctly identified after a median of only 15 element presentations but the pattern 1-3-4-2-6-4 required a median of 25 element presentations for correct identification. Furthermore, it is possible to identify patterns which are easy to identify by positions but hard to identify by directions (e.g., 1-5-3-6-2-4) and vice versa (e.g., 1-3-2-4-5-6). In the former case the median number of element presentations was 23 for the positional modes of responding and 16 for the directional mode of responding, in the latter case the corresponding data were 16 and 24 element presentations. Continuing analysis of these data will examine structural properties of the objective patterns to determine the relative efficacy of different types of element features (e.g., position and direction) and different pattern components (e.g., runs, thrills, alternations) in predicting pattern perception.

Conclusion

The ultimate aim of Experiment 2 is to classify the 720 six-position patterns as a function of relative difficulty and the strength of alternative psychological organizations. The classification will serve as baseline data that will in turn allow studies like that described in Experiment 1 to be conducted with much greater control over relevant features of the perceptual and motor task. It is proposed that these future studies will show that the acquisition, retention, and retrieval of information concerning the structure of temporal patterns are mutually influenced by what the subject sees and how he responds to what he sees.

REFERENCES

Freides, D. *Psychological Bulletin*, 1974, 81, 284-310.

Gibson, E.J. *Principles of Perceptual Learning and Development.* NY: Appleton-Century-Crofts, 1969.

Handel S. & Buffardi, L. *Quarterly Journal of Experimental Psychology*, 1969, 21, 256-266.

McGee, D.H. & Christ, R.E. *Journal of Experimental Psychology*, 1971.

Pew, R.W. In B. Kantowitz (ed), *Human Information Processing: Tutorials in Performance and Cognition*. Potomac, Md.: Lawrence Erlbaum Assoc., 1974.

Royer, F.L. & Garner, W.R. *Perception & Psychophysics*, 1966, 1, 41-47.

BEHAVIORAL SCIENCE ANALYSIS OF DOD FOOD SERVICE: IS CONSUMER ACCEPTANCE ENOUGH?

Harry L. Jacobs
U.S. Army Natick Research and Development Command

The current DoD R&D program in food service maximizes *consumer acceptance*, measured by psychologists, and used by developers. The goal is to provide highly acceptable food to *maximize consumption* and *minimize waste*. Unfortunately, food intake and body weight in the military mirror our overfed, underexercized society. Preference is for the high calorie, high fat diet nationally correlated with cardiovascular disease, a problem recognized in the civilian sector by recent Congressional discussions of *dietary goals*, and in the military by new emphasis on physical fitness, without reference to food intake. When Command rethinking recognizes that food service is part of the cause and could be part of the solution to this problem, psychologists can make a significant contribution by starting to look at the military *consumer* as a *member of the Armed Forces*, relating food acceptance to the individual's *behavioral effectiveness*, and initiating projects to modify food habits.

The Natick Laboratories are responsible for the application of Food Science and Food Technology to the development and evaluation of military food and food service systems for all of the four military services. The central concept in Behavioral Sciences support for this DoD Food Program is *consumer acceptance* (Figure 1). Our psychologists use *food preference surveys to provide data on food acceptability* for product development and menu planning groups (Figure 1, left) and *consumer opinion surveys* to provide data on various aspects of system acceptability for those responsible for system design and evaluation (Figure 1, right).

Figure 2 lists the major factors in dining hall satisfaction that we find in consumer opinion surveys. Although we usually find food quality variables topping the list, there are cases where food service variables become critical, e.g., where individuals have long waits in line. Figure 3 shows clusters of *high preference* and *low preference* food items in large sample food preference surveys. These were obtained from an aggregate of 4,000 men at 12 test sites with Army, Air Force, Navy, or Marine garrison feeding systems.

This approach has produced a very successful interdisciplinary partnership between the psychologist, the system developers at Natick, and the military representatives for the four services that are responsible for the successful operation of these systems in the field. It is an almost ideal example of the operation of human factors analysis within a prescribed system setting. The psychologist gets the opportunity to evaluate food items and food service equipment in the laboratory context, early in the development cycle at Natick, as well as the opportunity to evaluate the same items as they operate in actual garrison and field feeding systems.

Over the past decade, the DoD food program has been increasingly successful in developing increasingly efficient and increasingly acceptable food service systems. Figure 4 summarizes the current long range goals of the DoD food program from this point of view. Food scientists and technologists focus on food service goals, to maximize (1) shelf life, (2) microbiological and nutritional stability, (3) convenience, in product development and distribution logistics, in storage, preparation, cooking, serving and waste disposal, and (4) cost effectiveness, throughout the whole system. The behavioral scientists focus on consumer goals, to (1) increase acceptance, (2) increase attendance, (3) decrease food waste, and (4) increase morale in the dining facilities.

Since the military does not have the profit indicators of success available to industry, we consider the job complete when we provide the consumer with good, nutritious food, what he wants, and as much as he wants, with maximum consumption and minimum waste (Figure 4). Thus, we are really interested in *foods*, not *people*. This is reflected in the fact that the food program itself is managed as part of the logistics system in DoD. Thus, providing food for troops is much like providing gasoline for military vehicles.

In the case of military vehicles, consumption parameters are built into the engine design, based upon work expenditure. There is no question of whether the truck "likes" the gasoline, or whether it will ingest more or less than it needs. Unfortunately, this analogy does not apply to modern man in a sedentary society where national lifestyle produces an increasing proportion of overfed, overweight, and sedentary citizens, and where the major public health problem is death and disability from cardiovascular diseases related to food habits

Over the past year, Congressional discussion of these well-known facts has produced a national debate on dietary goals. (Jacobs, 1978). Some of this type of thinking is beginning to be reflected in the military, e.g., over the past year, a new Army regulation on physical fitness and weight control was initiated, and is being personnally monitored by the Army Chief of Staff. This regulation has produced a lot of activity in the leadership responsible for military operations. There is stronger command control in exercise programs and real threat of discharge for failure to comply with body weight and physical fitness standards. At present, effort is focused on exercise and physical fitness, without mention of military food service and only cursory attention to food *intake* as a causal factor.

If one looks at military food service from the point of view of food habits, weight control, and dietary goals, rather than from the logistically oriented model shown in Figure 4, an interesting picture appears (Jacobs and Meiselman, 1976). A clear look at the mass of data we have collected on food preferences shows the same pattern found in the civilian sector. Looking at Figure 3 again, it is apparent that high fat, high calorie items are preferred. Short order snack food items are the most popular of the highly-liked items, whole milk is preferred to skimmed milk, ice cream to yogurt, and cola to lo-cal soda. All available per capita consumption data also shows that daily food intake is consistently high in calories and fat.

In one study of 100 troops monitored in detail in the military for 28 days, the troops gained an average of 4.4 pounds by overeating as much as 1000 calories per day on highly palatable rations (Figure 5). Analyses of fat intake showed that mild ingestion provided 40% of the fat calories (Figure 6). If we could shift acceptance from whole to skimmed milk, this would significantly decrease fat intake. This untapped area of food habits research provides unlimited opportunities for behavioral scientists.

What of the physical fitness of these troops? Figure 7 shows that the smaller, lighter men run faster, have a stronger grip, and a better circulatory response to exercise. When one considers anthropometric data collected by Natick staff over the years on large samples of servicemen, it is apparent that the average serviceman is 4-5 pounds heavier and about an inch larger in the waist (Figure 8). This increase in weight is independent of age, and has recently been verified in the civilian sector by the National Center for Health Statistics. Thus, it is clear that body weight is increasing in the military and is related to physical performance. Since food habits combine with exercise (or lack of it) to determine body weight, there is now an opportunity for the behavioral scientist to begin to consider the food service system, both as possible cause, and as possible solution to problems of weight control, physical fitness, and dietary goals.

This requires a new orientation in which the serviceman is considered in a dual role as a *member of the Armed Forces* (Figure 9) in addition to the role of *consumer* to which we have limited him in our program in the past (Figure 4). Thus, in Figure 9 we expand our classical question of whether food service is related to the dining hall alone, asking now whether it applies to the morale and general performance of the consumer as a member of the Armed Forces, doing his assigned job.

Figure 10 outlines a new model for behavioral science analysis of the effect of food service systems on the individual, from this point of view. Starting from the fact that food acceptance determines choice and intake, we can ask the general question of whether the pattern of food habits in our overfed, underexercised serviceman affects him as an individual. Most of the public discussion on dietary goals is on the mass of evidence on food habits and *health* (left of Figure 10) focusing on the strong correlation between high calorie, high fat diets, and overweight as risk factors in cardiovascular disease. Unfortunately, knowledge of these correlations has not been very effective in changing life-style in this country. Perhaps it is because of the long-term character of the health effects, which usually appear later in life that it is the most difficult to initiate corrective action in (as in the case of smoking and cancer).

In the case of biologically defined *performance*, (center of Figure 10), we know the parameters relating physical fitness and body weight (e.g., Figure 7). This type of data is already influential in DoD actions, as in the new Army regulation on physical fitness discussed above. On the right of Figure 10, we have used the term *behavioral effectiveness* to cover the effects on *psychobehavioral* performance that we know least about, and where behavioral scientists can make the greatest contribution. This includes attention, vigilance, motivation, psychomotor skills, language skills, problem solving ability, etc., all critical factors in defining individual competence in carrying out military tasks.

For decades, conditions of sedentary life in the military have continued to grow as we have substituted advanced technology for labor intensive duties. In recent years, budget tightening has increased pressure to minimize manpower and costs. Analysis of the role of food service and food habits on behavioral effectiveness of the Armed Forces, and development of techniques to modify food acceptance and food habits themselves could pay major dividends in understanding and improving many operational systems in the Armed Forces in the future.

REFERENCES

Jacobs, H. L., National ferment in dietary goals: Challenge and opportunity for the Defense Department's food program. *Activities Report, Research and Development Associates*, 1978, 30 (In Press).

Jacobs, H.L. & Meiselman, H. L., Customer morale and behavioral effectiveness: Accomplishments and goals of psychological studies of food service systems. *U.S. Army Natick Research and Development Command Technical Report*

PROJECT DEVELOPMENT
MENU PLANNING

SYSTEM DESIGN
EVALUATION

"CONSUMER ACCEPTANCE"

FOOD PREFERNCE SURVEYS

BEHAVIORAL SCIENCES
SUPPORT ACTIVITIES

CONSUMER OPINION SURVEYS

Figure 1.
Consumer Acceptance In DoD Food Program

FACTORS IN CONSUMER OPINION OF FOOD SERVICE SYSTEMS

FOOD	FOOD SERVICE	ENVIRONMENT	SOCIAL
QUALITY	PERSONNEL	CONVENIENCE	EATING COMPANIONS
QUANTITY	SPEED OF SERVICE	PHYSICAL	ATMOSPHERE
VARIETY	HOURS OF OPERATION	DECOR	
	EXPENSE	MONOTOMY	

Figure 2.
Typical Results of Consumer Opinion Surveys of DoD Food Service.

FOOD CLASS	HIGH	LOW
FISH AND SEAFOOD	FRENCH FRIED SHRIMP	BAKED FISH
	SEAFOOD PLATTER	SALMON
	LOBSTER	BAKED TUNA & NOODLES
MEATS	ROAST BEEF	GRILLED LAMB CHOPS
	SWISS STEAK	SPARERIBS W/SAUERKROUT
	POT ROAST	CORNED BEEF
	GRILLED STEAK	PORK HOCKS
	GRILLED MIN. STEAK	PICKLED PIGS FEET
	B B Q SPARERIBS	SAUERBRATEN
	GRILLED HAM	
	BAKED HAM	
	ITALIAN SAUSAGE	
	FRIED CHICKEN	
	BAKED CHICKEN	
	HOT TURKEY SAND. W/GRAVY	
	HOT ROAST BEEF SAND. W/GRAVY	
STEWS & EXTENDED MEATS	LASAGNE	CHICKEN CACCIATORE
	PIZZA	CHILI MACARONI
	SPAGHETTI W/MEAT SAUCE	HAM LOAF
	SPAGHETTI & MEAT BALLS	VEALBURGER
	MEATLOAF	STUFFED CABBAGE
	SWEDISH MEATBALLS	CORN BEEF HASH
	SALISBURY STEAK	STUFFED GREEN PEPPERS
	BEEF STEW	PORK CHOP SUEY
		SWEET & SOUR PORK
		SUKYAKI
		BAKED TUNA & NOODLES
SHORT ORDER SANDWICHES	HAMBURGER	FRANK, CHEESE & BACON
	CHEESEBURGER	SALAMI SAND.
	HAM SANDWICH	BOLOGNA SAND.
	BLT	HOT REUBEN SAND.
	GRILLED CHEESE	HOT PASTRAMI
	GRILLED HAM & CHEESE	FISHWICH
	SLOPPY JOE	
	PIZZA	
FRESH FRUIT	ORANGES	PLUMS
	APPLES	HONEYDEW MELON
		FRUIT CUP
MILK PRODUCTS	MILK	SKIMMED MILK
	ICE CREAM	BUTTERMILK
		FRUIT-FLVD. YOGURT
CARBONATED BEVERAGES	COKE	LO-CAL SODA

Figure 3.
High and Low Consumer Acceptance in Armed Forces Food Preferences Surveys.

CURRENT PROGRAM

"HERE IS GOOD FOOD - EAT WHAT YOU LIKE AS MUCH AS YOU LIKE"

FOOD SERVICE	CONSUMER
SHELF LIFE	INCREASE ACCEPTANCE
STABLE	INCREASE ATTENDANCE
CONVENIENT	DECREASE WASTE
COST EFFECTIVE	INCREASE MORALE
ETC.	ETC.

Figure 4.
Model of Goals of Current Overall DoD Food Program

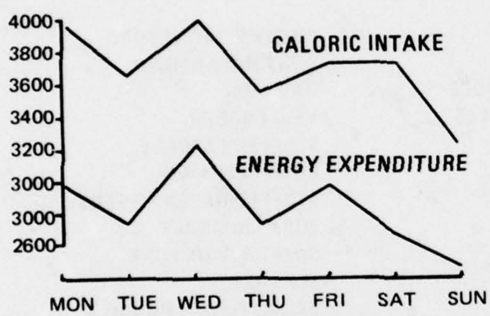


Figure 5.
Typical "Overeating" Pattern By Moderately Active Soldiers On Highly Acceptable Rations.

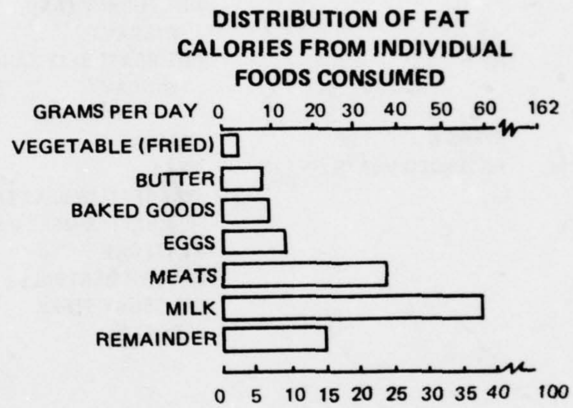


Figure 6.
Milk—Highly Preferred Source Of High Fat Intake In Troops.

MEASURE	AVERAGE FOR BEST 9	AVERAGE FOR OVERALL GROUP
FINAL WT. (LB.)	152.35	159.19
HEIGHT (IN.)	57.41	69.9
MEAN OF FIVE SKINFOLDS (CM.)	0.913	1.354
PERCENT BODY FAT	9.58	12.61
MNL STEP TEST	162.7	173.73
HAND DYNOMETER (LB.)	117.48	110.52
CONTEST MARCH (MIN.)	70.86	78.62
B.M.R. (CAL/HR.)	73.68	75.06

Figure 7.
Body Size and Physical Performance in Troops—An Inverse Correlation?

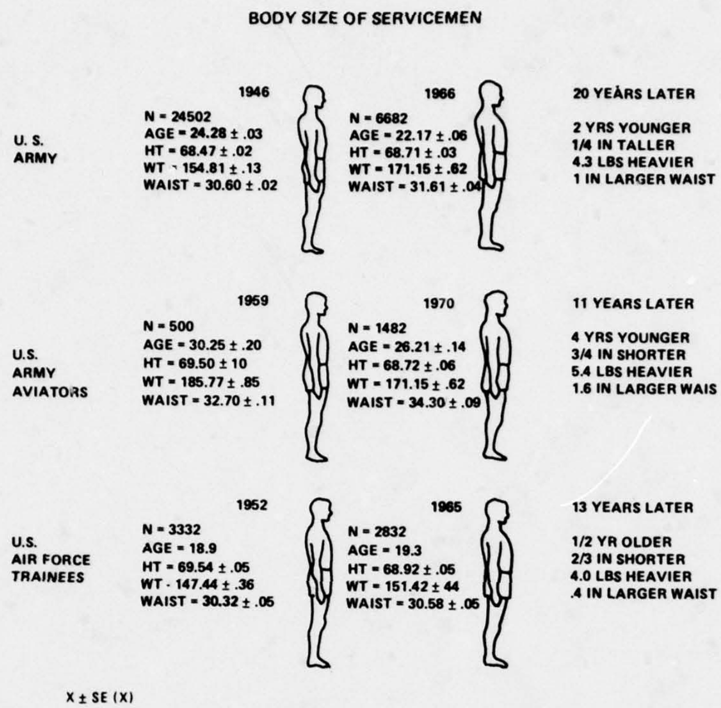


Figure 8.
Increased Body Weight in the Armed Forces Since World War II.

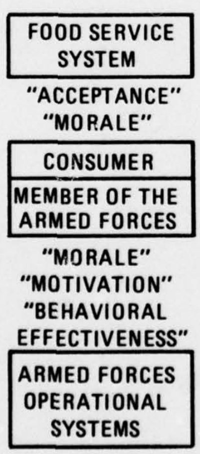


Figure 9.
Does Food Service Affect The Military Consumer In His Role As Soldier?

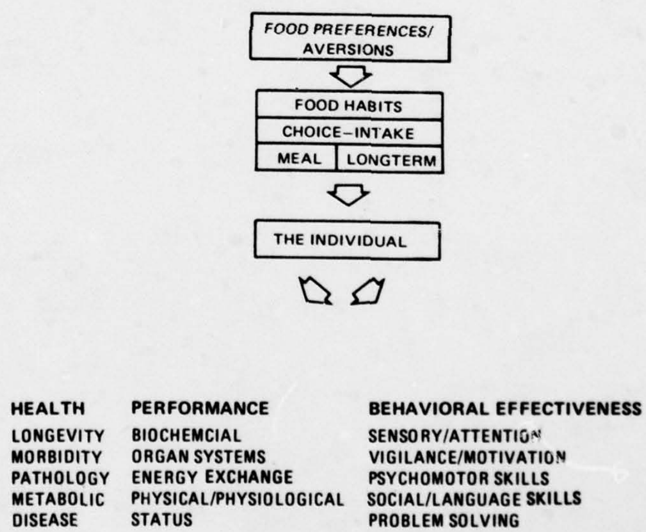


Figure 10.
Food Habits and Behavioral Effectiveness—An Untapped Opportunity For Behavioral Science.

TRANSFER, INFORMATIONAL FEEDBACK AND INSTRUCTIONAL SYSTEMS DEVELOPMENT

Charles W. Howard, Ph.D.
Army Research Institute

INTRODUCTION

The field of instructional technology includes research and development in the areas of programmed instruction, audio visual equipment and digital computers. Each area has developed instructional materials or software to improve the quality of instruction. One important problem is how to modify existing instructional programs which have not been developed utilizing a systems approach. In addition to insuring post-instructional competence of the trainee, instructional strategies must also be directed toward the transfer of learned skills to applications of those skills. A fundamental problem shared by instructional technologists is modifying or developing materials from existing content materials. Each area has developed instructional materials specified by objectives and evaluated by objective based tests to insure the quality of instruction. Instructional technologists utilize the Instructional Systems Development (ISD) approach for analyzing the performances to be learned and designing instructional systems to insure that the desired instructional outcomes are realized.

The Department of Army has begun converting its training program into self-instructional sets of materials. Numerous programs are currently self-instructional, however, many instructor-taught programs are still used and in need of adaptation. James E. Briggs (1964), Smith (1966) identified two potential problems in converting existing lecture demonstration course materials into self-instructional programs which include 1) time to produce effective sets of materials; 2) available programmers. Because of time constraints and a lack of available programmers, alternate strategies for converting typical course materials need to be analyzed.

With this mission, it is necessary to examine (a) the effects of alternative strategies for instructional development of self-instructional materials, and (b) whether instructional developers can modify existing software into self-instructional programs which promote learning by employing auto-elucidation techniques.

The primary purpose of this study was to determine the efficiency of five types of instructional development strategies. Efficiency was investigated in terms of achievement and time. The secondary purposes were 1) to investigate the utility of information feed-back for each instructional development strategy, 2) to investigate the utility of adjunct programming in the form of post text questions and remedial branching, and to explore the effects of post-instructional performance on a transfer test related to the learned skills.

The five instructional strategies that were explored in this study were:

1. instructional text
2. instructional text supplemented by post text questions;
3. instructional text supplemented by post text questions along with knowledge of results of the post text questions;
4. instructional text supplemented by post text questions, knowledge of post text questions and directions for remedial instruction;
5. programmed instruction, developed using the identical content of the instructional text and containing linear and branching programming.

The independent variables in this study were instructional strategies and informational feedback on the post test. The dependent variables in this study were transfer test scores, adjusted posttest scores , posttest time, total package time and total teaching time.

METHODOLOGY

Sample: The population from which the sample was drawn consisted of persons serving as enlisted or commission engineering personnel for the US Army. This study was conducted at Fort Belvoir Army Engineering School, Belvoir, VA.

The sample consisted of 720 subjects who were assigned to various technical training programs by the Department of the Army. The subjects were randomly assigned to one of the six groups, i.e., five treatment groups and one control group. This random assignment yielded 120 subjects per group.

Pre-test scores were used to determine the level of mastery of each subject. Each subject was classified as a non-master and therefore admissable as a subject for this study.

Treatment: The five instructional strategies consisted of content material, questions, knowledge of results and remediation appropriate to the strategy.

The primary component was the content material. The content material used for this study described the identification and functions of components for a low-voltage circuit tester.

The next component of the instructional strategy was the issuance of statements of questions. The questions were derived from the content material and were designed to facilitate learning for knowledge and comprehension of the content material.

Knowledge of results (KR) was the next component. This component consisted of subject matter experts responses to the questions. For the puproses of the study, distinctions between instructional strategies were made by providing KR to the learner in three of the five instructional strategies. The inclusion of KR was a form of feedback, however the implied use of KR was limited to informational feedback. Thr KR component was designed to provide the learner with the model performance expected and therefore enable the learner to make a comparison of the model performance and his actual performance.

The final component of instructional strategy was remediation. This component consisted of a written statement directing the subject to appropriate portions of the content material when actual performance made by the subject was not in agreement with the model performance.

Design: The research design proposed to test the hypotheses for this study consisted of two analysis of variance models and one analysis

of covariance model. The first model was designed to test the dependent variable transfer test score. The two factors were instructional strategy (six levels) and feedback (two levels). The second model was designed to separately test the dependent variables: posttest time, total package time, and teaching time. The factor for model two was instructional strategies (six levels). Model three was designed to test the dependent variable posttest score adjusted for the influence of the pretest score. The independent variable was instructional strategy.

The F ratios resulting from the models for each main effect and interaction was tested with the probability to falsely refecting a hypothesis set at 5%. The procedures employed to test pairwise comparisons of means for significant F ratios were Student, Newman-Keuls and Duncan simultaneous significant post-hoc tests.

Procedure: The administration of each of the five instructional strategies followed a common procedure consisting of five phases.

Phase I. Distributed a set of instructional objectives to each subject of each treatment group, as well the subjects in the control group.

Phase II. The subjects were given the package pretest for a maximum period of 30 minutes. The subjects' length of time for testing was recorded on his individual test by the monitor for analysis in this study.

Phase III. Subjects were given a set of instructional materials appropriate to their randomly assigned treatment group. The maximum length of instructional time was 50 minutes for each group. When a subject completed the materials he returned the materials to the monitor. The monitor recorded the amount of time for each subject.

Phase IV. Immediately following the treatment the subjects were given a posttest on the materials. The maximum length of time for the posttest was 30 minutes. The same procedure for recording the subjects' time for taking the posttest was used as was used in the pretest setting.

Phase V. Following the posttest, the subjects were randomly selected to either receive informational feedback on their posttest results or not to receive informational feedback.

Administration of Transfer Test: On the same day the instructional materials were administered a transfer test was administered. This test required an application of the cognitive skills presented in the instructional materials.

The actual formation of the five instructional strategies consisted of incorporating the four components: content, questions, Kr, and remediation into the instructional development strategies as described below.

1) Programmed Instruction: Subject was given a self-instructional set of materials that had been developed to move the subject through the material.

2) Text: Subject was given a manual that contains the content material with no programming of the content material.

3) Text/Questions: Subject was given a manual which were at the end of the material. The directions instructed the learner to overtly respond to the questions.

4) Test/Questions/Knowledge of Results: Subject was given a manual and a set of questions as above. This instructional strategy was designed to include the correct responses for the questions in addition to text/questions.

5) Text/Questions/KR/Remediation: Subject was given a text containing the content material, questions, KR, and directions to portions of the test to review non-mastered content material.

6) Control: Subject was pretested and post-tested with no instructional treatment.

Instruments: The pretest and posttest consisted of twenty-seven matching items. The format was to provide each subject with a list of statements and a graphic representation of the low-voltage circuit tester. The graphic included the components of the low-voltage circuit statement with the numeric in the graphic.

The transfer test consisted of twenty-one completion items. The transfer test used a graphic workbook to substitute for the actual equipment described in the items. The transfer test was designed to measure the transfer to skills/knowledge presented in the instructional materials to the application of these skills and knowledge. The format was to provide each subject with a written statement of the conditions and standards and a graphic representation of the equipment. The subject was then required to 1) indicate the appropriate hookup, 2) the range of readings, or 3) trouble shooting procedure.

Results: The means, standard deviations, and Ns for the five instructional strategies and the control group on the dependent variables total package time, post-test time, teaching time, pre-test score, post-test score and adjusted post-test score are reported in Table 1.

The dependent variables total package time, post-test time and teaching time were analyzed by an analysis of variance model (ANOVA). The results of these analyses indicated a significant difference existed between the groups; total package time, $F[.95]\ (4,595) = 59.457$, $p<.05$; post-test time, $F[.95]\ (5,719) = 20.68$, $p<.05$; teaching time, $F[.95]\ (4,595) = 67.51$, $p<.05$. The pairwise comparisons were analyzed for each of these dependents using the Student Newman Keul's (SNK) post-hoc simultaneous significant test. The results of the SNK tests are condensed and reported in Table 2. The results indicate the text and question group was not significantly higher than the programmed instruction group in the dependent variables post-test time, teaching time or total package time. Thus, an instructional strategy has been identified that is as efficient as programmed instruction.

Evaluation of the effectiveness of the instructional strategies was measured in terms of performance.

The dependent variables pre-test and post-test performance scores were analyzed by an analysis of co-variance model (ANCOVA). The result of this analysis indicated a significant difference existed between the adjusted post-test performance scores reported in Table 1 with an $F[.95]\ (5,713) = 53.59$, $p<.05$. A test of homogenity of regression was performed and no violation of the assumption was observed. The pairwise comparisons were analyzed using the Duncan post-hoc simultaneous significant test. The results are shown in Table 3 and indicate the text-plus-question group was not significantly different that the programmed instruction group. Thus, an instuctional strategy has been identified that is as effective as programmed instruction.

Evaluation of informational feedback and transfer. The dependent variables transfer test score was analyzed by a factorial analysis of variance model. The results of this analysis indicated a significant difference existing between the group means for the factor instructional strategy with an $F[.95]\ (5,708) = 20.46$, $p<.05$. The factor informational feedback was not significant with an $F[.95]\ (1,708) = .08$, $<.05$. The interaction was not significant with an $F[.95]\ (5,708) = 1.28$, $p<.05$. The examination of the pairwise comparisons of the instructional strategies was evaluated with the SNK post-hoc procedure. The results indicated transfer of skills and knowledge occurred as well as programmed instruction in three alternative instructional strategies. The three alternatives were text plus questions; text, questions plus feedback and text only.

Discussion: The findings suggest an alternative instructional strategy has been developed that is as effective and efficient as programmed instruction. Programmed instruction results from a rigorous application of the ISD procedures and is also extremely time-consuming and therefore expensive. The instructional strategy which has been found to be as effective and efficient as programmed instruction uses the principles of mathemaenic behavior and auto-elucidation techniques. The application of mathemagenic activities and auto-elucidation techniques have cost effective implications for the conversion of instructional materials to self-instructional materials.

Conclusions: This research has examined alternative instructional strategies, transfer of cognitive skills, auto-elucidation (e.g. effects of KCR, remediation, posttext questions), and informational feedback. This investigation has focused on the effectiveness of five types of instructional strategies and the utility of informational feedback.

The results support the application of auto-elucidation techniques in the form of posttext questioning. Text or written material followed by posttext questioning resulted in performance as effective as programmed instruction. Posttext questions are therefore regarded as aids to the learner which stimulate the learning environment.

Posttext questions supplementing but following the reading of written prose or text have been demonstrated as an effective mixture of instructional materials. For example, the treatments text, text plus questions plus feedback or text plus question plus feedback plus remeditation when uncontrolled was found to be less effective than PI or the use of text plus posttext questions.

Informational feedback was provided to 50 percent of the subjects in each of the six groups. This feedback was in the form of KCR upon the completion of the posttest. Informational feedback in the form of KCR was not found to significantly affect the performance of subjects on the transfer test.

In summary, the results support the following conclusions: 1) the use of posttext questions in the form of text and posttext questions was demonstrated to be the most effective alternative strategy; 2) informational feedback in the form of posttest KCR was not found to significantly affect the performance of subjects on the transfer test; 3) the application of auto-elucidation components in the form of knowledge of results and remediation when uncontrolled was less effective than the use of posttext questions only.

IMPLICATIONS AND RECOMMENDATIONS: The primary purpose of this investigation was to identify the efficiency of five types of instructional development strategies.

This study investigated the effects of applying auto-elucidation techniques which were developed by Sidney Pressey. E. Z. Roghkopf and L. Frase further studied the process of incorporating questions into written prose. The implications are that auto-elucidation techniques in the form of post-text questions create a learning situation that is both meaningful to the learner and desirable to the instructor and the institution. The meaningfulness of the learning situation has been mathemagenic behavior. The use of post-text questions requiring overt response without any form of feedback produced the most desirable results both in terms of time and performance for the alternative instructional strategies.

This implication has major impact in the psychological area of feedback for curriculum developers, instructional technologists and the field of psychology. Arguments against this implication would generally indicate that treatment groups receiving text + questions and feedback would presumably obtain higher scores. However, as demonstrated within this study, this was not the case. Therefore, instructional materials for which treatment groups do not have access to the correct answers but are permitted to review the instructional materials, are both effective and efficient.

Mathemagenic behavior as described in the literature and discussed in this study can be present in the absence of informational feedback.

The instructional strategies investigated have incorporated the principles of system analysis, informational feedback and auto-elucidation processes.

Instructional development models used in this study yielded instructional materials that were efficient in terms of reducing teaching time and effective in terms of performance. Therefore, the conversion of existing instructional materials into self-instructional modules using mathemagenic principles can have cost-benefit advantages for educational, industrial and Armed Service institutions.

The implications for cost-effective instructional materials development are as follows: 1) Material may be developed that yield performance scores higher than existing ISD developed materials while requiring less teaching time; 2) The cost associated with development of the materials may be significantly less.

The next implication concerns the Armed Services. The Armed Services have two plans under which all training operates. The two plans are peacetime and wartime. Without discussing either of the plans in detail, the fact remains that during wartime the Armed Services must reduce the training time of soldiers in order to meet the demands. Therefore, considerations must be made regarding the instructional materials. Results of this study indicate that the difference in the means for text (2) and question (3) groups total teaching time was 6.76 minutes which represents one-third less time. The decrease in performance was 3.37 test items or 12% of the total possible score. Without discussion about the future effectiveness of the soldier, the point to be considered is that alternative instructional strategies may cut training time by 1/3 therefore better meeting the wartime needs of our country.

The views expressed in this paper are those of the author and do not necessarily reflect the views of the US Army or the Department of Defense.

This paper is a portion of the author's thesis conducted at The Catholic University of America.

TABLE 1.
MEANS AND STANDARD DEVIATIONS OF TEACHING TIME, POST-TEST TIME, TOTAL PACKAGE TIME PRE-TEST SCORE, POST-TEST SCORE, AND ADJUSTED POST-TEST SCORES BY INSTRUCTIONAL STRATEGY

GROUPS		TEACHING TIME		POST-TEST TIME		TOTAL PACK TIME		PRE-TEST SCORE		POST TEST SCORE		ADJ.
	n	M	SD	M	SD	M	SD	M	SD	M	SD	M
PI	120	37.40	9.95	14.08	7.10	51.48	9.98	3.14	3.54	12.98	6.50	12.93
T	120	22.58	8.48	11.55	5.25	34.04	10.17	1.81	2.63	10.53	5.76	11.05
TQ	120	38.05	7.96	12.75	7.00	50.80	10.17	2.52	2.80	14.21	6.08	14.42
TOF	120	29.42	7.55	11.05	5.42	40.48	10.37	4.13	3.51	12.42	5.18	11.96
TOFR	120	33.04	8.24	10.67	6.04	43.72	11.22	2.78	2.61	10.54	6.70	10.55
CTRL	120	a	a	6.73	4.78	a	a	3.80	2.67	4.12	4.02	3.79

a = TREATMENT NOT ADMINISTERED TO CONTROL GROUP

TABLE 2.
PAIRWISE COMPARISON OF MEAN DIFFERENCES WITH FIVE INSTRUCTIONAL TREATMENTS ON TOTAL PACKAGE TIME

q(r, 595)		2.80	3.36	3.69	3.92
Ordered Group Mean	2	4	5	3	1
(2) Text	-	*	*	*	*
(4) T+Q+F		-	*	*	*
(5) T+Q+F+R			-	*	*
(3) Text+Quest.				-	ns
(1) PI					-

* = $p < .05$, $q(r, 595) \sqrt{MSe/n} = q(r,595)(.948)$

PAIRWISE COMPARISON OF MEAN DIFFERENCES WITH FIVE INSTRUCTIONAL TREATMENTS ON POST-TEST TIME AND THE CONTROL GROUP

q(r,595)		2.80	3.36	3.69	3.92	4.10
Ordered Group Mean	6	5	4	2	3	1
(6) Control	-	*	*	*	*	*
(5) T+Q+F+R		-	ns	ns	*	*
(4) T+Q+F			-	ns	ns	*
(2) Text)				-	ns	*
(3) T+Quest					-	*
(1) PI						-

* = $p < .05$, $q(r,714) \sqrt{MSe/n} = q(r,714)(.5475)$

PAIRWISE COMPARISON OF MEAN DIFFERENCES WITH FIVE INSTRUCTIONAL TREATMENTS ON TEACHING TIME

q(r,595)		2.80	3.36	3.69	3.92
Ordered Group Mean	2	4	5	1	3
(2) Text	-	*	*	*	*
(4) T+Q+F		-	*	*	*
(5) T+Q+F+R			-	*	*
(1) PI				-	ns
(3) Text+Quest					-

* = $p < .05$, $q(r,595) \sqrt{MSE/n} = q(r,595)(.774)$

TABLE 3.
PAIRWISE COMPARISON OF MEAN DIFFERENCES OF THE SIX GROUPS ON ADJUSTED POST-TEST SCORE AND THE CONTROL GROUP

			2.80	2.95	3.05	3.12	3.22
Ordered Group Mean		6	5	2	4	1	3
(6)	Control	–	*	*	*	*	*
(5)	T+Q+F+R		–	ns	ns	*	*
(2)	Text			–	ns	ns	*
(4)	T+Q+F				–	ns	*
(1)	PI					–	ns
(3)	T+Quest						–

* = $p < .05$, q(r,713 [**] = q(r,713)(.71238)

**NOTE: The adjusted MSe is an estimate of the average standard error (Dayton, 1970, pp. 322).

$$q = \frac{2MSerror^*}{n} \qquad \frac{SS\ treat}{1 + (p - 1)\ SSx\ error}$$

REFERENCES

Briggs, James E., Programmed Instruction: Break through in Air Force Training. *Trends in Programmed Instruction*, Ofiesh, Gabriel D., Meierhenry, Wesley C. (Ed.) 1964.

Frase, L. T., Learning from prose material: length of passage, knowledge of results, and position of questions. *Jounral of Educational Psychology* 1967, 58, 266-272.

Frase, L. T., Effects of questions, location, pacing, and mode upon retention of prose materials. *Journal of Educational Psychology*. 1968, 59, 244-249.

Frase, L. T., Questions as aids to reading: Some research and theory. *American Educational Research Journal*. 1968,5, 319-332.

Frase, L.T. Some data concerning the mathemagenic hypothesis. *American Educational Research Journal*. 1968, 5, 181-185.

Frase, L. T., Boundary conditions for mathemagenic behaviors. *Review of Educational Research*. 1970, 40, 337-347.

Pressey, S. L., Auto presentation versus auto-elucidation. *Programmed Instruction*. 2, 6-7, April 1963.

Pressey, S. L., Auto-elucidation without programming. *National Society of Applied Instruction Journal*. 3, 12-13, August 1964.

Pressey, S. L. and Kinzer, J. R., Auto-elucidation without programming, *Psychology in the Schools*. 1, 4, 1964, 359-365.

Rothkopf, E. Z., Learning from written materials: an exploration of the control of Inspection behavior by test-like events. *American Educational Research Journal*. 1966, 5, 86-91.

Rothkopf, E. Z. and Bibicos, E., Selective facilitative effects of interspersed questions on learning from written materials. *Journal of Educational Psychology*. 1967, 58,56-61.

Rothkopf, E. Z., Experiments on mathemagenic behavior and the technology of written instruction. Murry Hill, New Jersey. Bell Telephone Laboratories, 1968 (Mimeo.).

Rothkopf, E. Z., The concept of mathemagenic activities. *Review of Educational Research*. 40, 1, 325-336, 1970.

Rothkopf, E. Z., Variable adjunct question schedules, interpersonal interaction, and incidental learning from written material. *Journal of Educational Psychology*. 1972, 63, 87-92.

Smith, R. G., *The Design of Instructional Systems*. HumRRO Technical Report 66-18, 1966.

EMPIRICAL VALIDATION OF SELECTED INSTRUCTIONAL QUALITY INVENTORY PRESCRIPTIONS[1]

John A. Ellis Wallace H. Wulfeck II
Navy Personnel Research and Development Center

Norman D. Wood Meryl Baker M. David Merrill
Courseware, Inc.

Problem

The Navy is currently involved in systematically analyzing, redesigning, and redeveloping a large number of its technical training courses. To facilitate accomplishment of this task, the Instructional Quality Inventory (IQI), an instrument for identifying defects in instructional material and for prescribing revisions thereto, was developed under contract by the Navy Personnel Research and Development Center. During development several variables were identified that were hypothesized to affect the consistency and adequacy of instructional materials. However, many of the prescriptions relating to these variables have not been empirically validated.

Overview of the Instructional Quality Inventory

The IQI is an analytic tool that facilitates the evaluation and revision of existing instruction and the design of new instruction. It consists of a set of prescriptions that allows the user to inventory instructional quality and to prescribe revisions that will increase its effectiveness.

The IQI is designed to evaluate instructional materials on two main criteria: consistency and adequacy. Consistency must be determined before adequacy can be assessed.

The consistency criterion is met if it is determined that the instructional objectives, test items, and the instructional presentation are consistent. This is accomplished in two steps. First, the instructional objectives and test items are classified on two dimensions: (1) the performance, or task level, required of the student, and (2) the type of instructional content. These two dimensions then are combined to form a task content classification matrix, which is used to classify objectives, test items, and instructional presentations. This matrix is illustrated in Figure 1. If an objective and its corresponding test item can be classified in the same cell of the matrix, they are considered to be consistent. The second step involves rating the consistency between instruction and objective/test items. The IQI requires that different components of instructional presentation, called primary presentation forms, be present for different task levels. If the primary presentation forms required for the task level of each objective/test item is present, then the instruction is consistent with the objectives/test items.

Once it has been determined that instructional materials are consistent, the adequacy criterion is assessed. This is done by determining whether or not the instructional presentation adequately communicates the "to-be-learned" information. Variables that are to hypothesized as affecting instructional adequacy during IQI development include the following: (1) isolation (i.e., is the relevant information separated and clearly identified?), (2) help (i.e., is explanatory or mnemonic information provided?), (3) matching (i.e., are the examples and practice items matched, and (4) sampling (i.e., are the examples and practice items divergent). Instruction is evaluated on these variables to obtain an adequacy index. Each primary presentation form within the instruction may be rated as more or less adequate.

In the following sections, the task/content matrix will be described in greater detail, the primary presentation forms will be discussed, and validation experiments testing the consistency and adequacy assumptions will be summarized.

Task/ Content Classification Matrix

As shown in Figure 1, the task demension of the task content classification matrix is comprised of two major levels, *Use* and *Remember*. *Use* is defined as the act of applying a general relationship or definition to a specific situation where it has not been previously applied; and *remember*, as the act of bringing to mind something that has been seen before. Thus, a use item (or objective) would require the student to respond by applying a definition to a new example: that is, one that has not been seen previously as part of the instructional presentation. A remember test item (or objective) would require the student to respond by recognizing or recalling a definition or example that has been seen before. *Definition* is defined as a statement of relationship that can be applied to more than one specific object or event; and *example* is defined as a specific objective or event or its representation that does or could exist in the real world.

The use level cannot be divided—it always requires *new* examples. The reason for this is obvious—if an example had been previously seen, the test item or objective would be classified at the remember level. The remember level, however, can be divided into two sublevels depending upon what type of information the student is required to remember. Thus, *remember* definition means that the *remember example* means that the student must recall a previously seen example.

As shown in Figure 1, the content dimension of the matrix involves four mutually exclusive content categories: facts, concepts, procedures, and principles. Except for facts, for which there can be no definitions, all types can be tested at any of the task levels. These categories are defined as follows:

1. A *fact* is a one-to-one association of a symbol or name and a specific object or event.

2. A *concept* is a class of objects, events, or symbols that (a) share critical attributes, (b) can be referenced by a name or symbol, and (c) have discriminably different individual members.

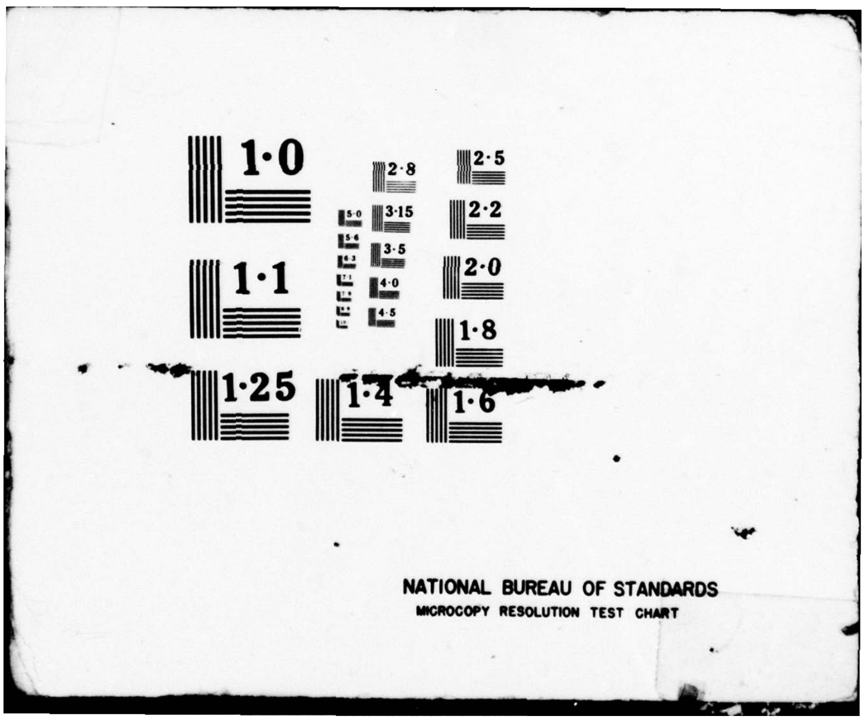
[1]The views and opinions expressed herein are those of the authors, and do not necessarily represent official policy of the Department of the Navy or the Department of Defense.

AD-A068 221 AIR FORCE ACADEMY COLO F/G 5/10
PROCEEDINGS OF THE ANNUAL SYMPOSIUM ON PSYCHOLOGY IN THE DEPART--ETC(U)
APR 78 J M BERMUDEZ, W H CLOVER, L J DAHLE

UNCLASSIFIED USAFA-TR-76-12 NL

4 OF 4
ADA
068221

END
DATE
FILMED
6-79
DDC


1·0
2·8
2·5
3·15
2·2
3·5
2·0
1·1
4·0
4·5
1·8
1·25
1·4
1·6
NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

TASK LEVEL	CONTENT TYPE			
	FACT	CONCEPT	PROCEDURE	PRINCIPLE
USE Definition on new examples				
REMEMBER Definition				
REMEMBER Example				

Note: Facts can only be tested at the Remember Example level.

FIGURE 1.
IQI TASK CONTENT CLASSIFICATION MATRIX.

3. A *procedure* is a series of steps required to produce an example of an outcome class. Each step may involve the temporal or branching decision, based either on a fact or the classification of an example of a concept. A procedure is often characterized as "how to do something."

4. A *principle* is a predictive relationship between specific examples of a concept, or among a set of related concepts, which explains why an example of a particular class is produced as a result of a particular manipulation.

Primary Presentation Forms

The IQI defines the instructional presentation form or display as the fundamental unit of instructional strategy. As indicated previously, the instructional presentation forms must meet consistency and adequacy requirements.

Four primary presentation forms or displays, which represent the various ways that information can be presented, have been defined:

1. Tell a definition (TD)—A display that presents a description of a concept, an algorithm that describes a proceudre, or a proposition that expresses a principle.

2. Tell an example (TE)—A display that illustrates how a definition applies to a specific example.

3. Ask about a definition (AD)—Not used in this research study.

4. Ask about an example (AE)—A display that presents an example and requires the student to respond to the example *or* it presents a name or definition and requires the student to respond by providing an example.

Objective

The overall objective of this study was to empirically validate some of the consistency and adequacy prescriptions of the IQI. Specifically, it was designed to evaluate six hypotheses concerning the consistency and adequacy of instructional materials.

Approach

Subjects participating in the study were enlisted men waiting to begin "A" School at the Propulsion Engineering (PE) School, Great Lakes. Instructional materials developed were based on PE School curriculum but were adapted to provide for 12 experimental treatments needed to test the six hypothehes. Four of these treatments represented *remember* level instruction; and eight, *use* level instruction. After students finished their instruction, they were on remember level test items (labeling and listing) and use level test items (classification). All subjects had the same testing materials.

Three experiments were conducted. Experiment I tested the consistency hypothesis, which stated that performance will decrease if test items and presentation strategies are not consistent. Experiment II tested the adequacy hypotheses for remember level items, which stated that performance will increase with the use of both a mnemonic and a several-page distributed practice vs. one-page massed practice groups. Finally, Experiment III teeted the adequacy hypotheses for use level items, which stated that performance will increase with the use of isolated definitions, divergent examples, and attribute isolation elaboration, respectively. This was done by comparing the performance of (1) the isolated definition vs embedded definition groups, (2) the divergent example vs. convergent example groups, and (3) the attribute isolation vs. no attribute isolation elaboration groups.

Results

1. Experiment I—The students in the use level treatment groups scored significantly higher on use level items (classification) than those in the remember level groups, and students in the remember level groups scored significantly higher on remember level items (labeling and listing) than those in the use level groups. Thus, the consistency prescription of the IQI is supported.

2. Experiment II—There were no significant differences in performance on the remember level test items for either the mnemonic or practice variable, but there was a significant savings of time for the massed practice condition. Thus, the two presentation adequacy prescriptions of the IQI for remember level items were not supported.

3. Experiment III—Students in the isolation treatments scored higher on all performance measures and took less time than students in the embedded definition treatment. Students in the divergent example treatments scored higher on use test items than students in the convergent example treatment. There were no differences between students in attribute isolation and no attribute isolation elaboration treatments. Thus, only two of the three presentation adequacy prescriptions of the IQI for use level items were supported.

Conclusions

Although the results of this study did not support all of the IQI prescriptions, they did provide considerable evidence that the IQI is a valid instrument for predicting student performance and for evaluating the effectiveness of existing and newly developed instructional materials.

DEVELOPMENT AND APPLICATION OF A TAXONOMY OF TACTICAL FLYING SKILLS

Robert P. Meyer and Bernell J. Edwards
Air Force Human Resources Laboratory
Williams AFB, Arizona

PROBLEM

The perceptual-motor aspects of flying tasks are not amenable to task analysis procedures described in current Air Force instructional systems development literature. Until recently, there has been little guidance available to those attempting to conduct task analyses specific to the perceptual motor domain of flying skills. Further, it has not been clear how a descriptive analysis technique for flying tasks, if available, could be used in generating training strategies and programs.

The intent of the present research was to develop a data base and methodology of sufficient depth and flexibility to permit the eventual generation of a behavior based tactical flying training program. A descriptive task analysis was seen as a necessary but insufficient condition for a useful data base. It was also determined that the approach must be behavioral, rather than systems oriented, in order to yield useful data outcomes. The objective, then, was to produce an analytic tool to identify, define, and compare varieties of skill components within and among tactical flying tasks. Analysis at this level of detail implies a comprehensive classification system or taxonomy.

METHOD

The approach taken in this study was an outgrowth of a taxonomic project undertaken for undergraduate pilot training tasks several years ago. The basic procedures used in developing the tactical flying task area followed precedents from the earlier study. A literature review was conducted for the earlier study which included literature on general taxonomies. A list of these references is attached to this paper.

Since it was beyond the scope of this study to address the entire task domain of tactical flying for all tactical aircraft, a representative aircraft and representative tasks were selected. A sample of sixteen tactical maneuvers were selected: seven air-to-ground maneuvers and nine air-to-air maneuvers. The F4-E was selected since this aircraft has good air-to-air and air-to-ground capabilities. Also, there has been extensive pilot experience in the F4-E.

Development began with the generation of a surface level task analysis ormat. The surface analysis was a way of describing in behavioral terms each task element at a high level of detail. The analysis was defined as a sequence of cues, mental actions, and motor actions occurring in temporal proximity throughout the maneuver. These events described essential pilot/aircraft relationships in terms of a stimulus-organism-response (SOR) behavioral model. Terminology used in the taxonomy has tended toward a flying vernacular in the interest of future taxonomic applications.

Figure 1 shows the surface analysis process. The cues were the environmental stimuli perceived by the pilot throughout the task sequence. Cues were dilineated as the visual, aural, tactual, and kinesthetic perceptual elements. As perceived by the pilot, cues resulted in various types of cognitive processes which were termed mental actions. This was pragmatic rather than theoretical, since mental processing for purposes of a behavioral taxonomy was regarded as an input/output system rather than a psychological construct. Mental action descriptors were: discerns, the short term processing of specific cues; sustains, the continuous cyclic processing of cues over a given time; anticipates, the recall processing associated with long term memory used prior to executing a specific task element; and determines, the processing of multiple cues and associated decision making. The third surface analysis category was termed motor action. It encompassed the actual flying and system functions of the task. Descriptors in this category included, among others: coordinates, the movement by the pilot of two or more aircraft controls simultaneously; adjusts, the incremental regulation of a specific control; and activates, the discrete engagement of a specific toggle switch, push button, or similar control. The basic descriptors used in the surface analysis, thus, comprised a unique way of viewing the detail and complexity of behaviorally oriented flying tasks.

Basic data from which the taxonomy could be developed were required. Interviews were conducted with 35 experienced F4-E pilots and weapons system officers to obtain authentic, reliable data on the sixteen selected maneuvers. During each interview, individual maneuvers were discussed and documented in detail with notes, diagrams, and recordings. Researchers then generated descriptive specifications and drawings for all maneuvers using this interview data. The surface task analyses generated by the researchers were then reviewed by the pilots during follow-up interviews to verify that maneuver descriptions accurately reflected the original interview data. Revisions and corrections were then made and the analysis data finalized.

Surface analysis techniques were able to reduce flying tasks into definable sequences. A task sequence consisted of both outside and inside cues, a mental action which solved for that set of cues and resultant motor actions based on the mental action. Each sequence has a temporal proximity to the previous and succeeding task sequence. The cue, mental action, and motor action elements thus contained the basic skill information of each sequence. This sequencing, then, is analogous to the single frame of a motion picture which, when put together with other frames, produces the task action. The task sequences of the analyses were found to be the basic skill level of this research and the fundamental building blocks of the taxonomic data base.

The next step was to process the surface analysis into useable raw data for the taxonomic classification system. A set of classification rules using the alpha-numeric coding system, shown in Figure 2, was designed for this step. A notation card for each element sequence was cued to reduce behavioral data from the surface analyses into a number of skill categories to facilitate taxonomic classification. Cues were coded, tabulated and classified by type and quantity. An input index was developed to reflect a value ratio between the type and quantity of cues. Mental actions were coded and classified on each skill card to show the type of information processing or decision processing involved. Motor actions were coded and annotated according to continuity and type of motor output. An output index was also developed to reflect a ratio between the type and continuity of motor output responses. The main function of the skill and notation system was to provide a useful method for comparing the type, quantity, and complexity of the various behavioral elements once all that data had been assimilated into the hierarchical structure of the taxonomy.

A classification hierarchy matrix was developed to permit the sorting of notation cards into skill groups which exhibited the same or similar behavioral characteristics. The hierarchy was organized following the same nine descriptive categories used in the classification rules. These reflected the various combinations possible between the types and quantities of cue, mental action, and motor action categories.

For example, the placement of a specific notation card within the classification matrix was first determined by the simple dichotomy as to whether it contained a simple or complex mental action. Figure 3 shows the complete data system of the taxonomy. When the coded behavioral data had been entered on each notation card and the card had been assigned to the appropriate sorting slot within the classification matrix, the essential structure of the tactical task taxonomy was complete.

RESULTS

The taxonomy permitted a variety of applications. The variety and complexity of skills across various maneuvers were compared. The frequency of specific skills required for various selected maneuvers was compared and common elements identified. With this capability, the grouping or sequencing of basic skills was identified. For example, the taxonomic system has shown that, in general, air-to-air skills contain a proportionally large complement of air-to-ground skills. Conversely, air-to-ground skills, in general, contain only a scattering of air-to-air skills. Skill comparisons across tasks have important implications for the structure of training sequences. The across task basic skill comparisons, made possible through the use of the taxonomy, permitted the identification of skill criticality for various training objectives. It was also possible to analyze the skills within a given group of tasks in order to design a single standard task which would contain a high percentage of skills identical to those of the given task group. In this way the taxonomy could provide a shortcut to the development of a wide range of flying proficiencies for many maneuvers. The potentials for training cost benefits are obvious.

The surface task analysis was also used to determine the tentative functional requirements for task specific simulator capabilities. These requirements have been derived using the taxonomy with its limited sixteen-maneuver data base. Requirements were derived using an extensive cues analysis. However, it must be kept in mind that functional requirements for the total domain of tactical flying would require a complete taxonomic data base in addition to consideration of a range of related factors, including instructor subjective opinion.

Another application of the taxonomy has been the generation of a task/skill difficulty index. This index was based upon a weighting system across eight skill factors identified in the classification system. Applying the weighting formula across each task sequence for each maneuver yielded a cumulative difficulty index for each of the sixteen maneuvers. This permitted a ranking of the maneuvers according to overall difficulty as a recommended training order. The derived rank order has been generally confirmed through an independent subjective ranking of maneuvers by experienced TAC pilots. The ordering of maneuvers to be taught implies a building block approach to tactical training syllabus sequences.

A large variety of other potential applications exists for the taxonomy. The applications of taxonomy based flying training and skill maintenance programs are anticipated across many areas within the Air Force training system.

REFERENCES

Angell, D. et al, *Study of Training Performance Evaluation Techniques*. Naval Training Device Center NAVTRADEVCEN TR-1449-1, Port Washington, New York, October, 1964.

Brecke, F. and Gerlach, V., *Model and Procedures for an Objective Maneuver Analysis*. AFOSR-TR-73-0523 Technical Report No. 21201, Arizona State University, December, 1972.

Brecke, F. and Reiser, R., *Critical Components of Flight Instruction as Perceived by IPs and SPs*. Technical Report 21129 Grant No. AFOSR 71-2128, Instructional Resources Laboratory, Arizona State University, Tempe, Arizona, November 1972.

Chambers, A.N., *Development of a Taxonomy of Human Performance: A Heuristic Model for the Development of Classification Systems*. American Institutes for Research, Silver Spring, Md. Technical Report 4, March, 1969.

Cotterman, T. E. and Stolurow, L. M., *A Taxonomy of Learning Task Characteristics*. AMRL Technical Report TDR-64-2, Jan., 1964.

DeGreene, K. B., (Ed.) *Systems Psychology*. McGraw-Hill Book Company, New York, New York, 1970.

Earle, T. C., *Interpersonal Learning*. Oregon Research Institute Technical Report, Office of Naval Research Contract No. N00014-68-C-0431, Contract Authority Identification Number NR 153-311, May, 1971.

Federal Aviation Agency. *Flight Training Handbook*, No. AC 61-21, 1965.

Folley, J. D., *Development of an Improved Method of Task Analysis and Beginnings of a Theory of Training*. Naval Training Device Center Technical Report, NAVTRADEVCEN 1218-1, Port Washington, New York, June, 1964.

Gerlach, V. et al, *The Generation of Cues Based on a Maneuver Analysis*. Technical Report 21202, Grant No. AFOSR 71-2128, Instructional Resources Laboratory, Arizona State University, Tempe, Arizona, December, 1972.

Matheny, W. G., Lowes, A. L., Baker, G., and Bynum, J. A., *An Investigation of Visual, Aural, Motion, and Control Movement Cues*. Naval Training Device Center Technical Report NAVTRADEVCEN 69-C-0304-1, April, 1971.

Miller, R. B. "Task Taxonomy: Science or Technology?", *Ergonomics* 10 (2) pp. 167-176, 1967.

Rigney, J. W., "Maintainability: Psychological Factors in Design" in DeGreene, K. B. (Ed.) *Systems Psychology*. McGraw-Hill Book Company, New York, New York, 1970.

Smith, B. J., *Task Analysis Methods Compared for Application to Training Equipment Development*. Naval Training Device Center Technical Report NAVTRADEVCEN 1218-5, Port Washington, New York, September, 1965.

Staats, A.W. and Staats, C.K., *Complex Human Behavior, A Systematic Extension of Learning Principles*. Holt, Rinehart and Winston, New York, New York, 1964.

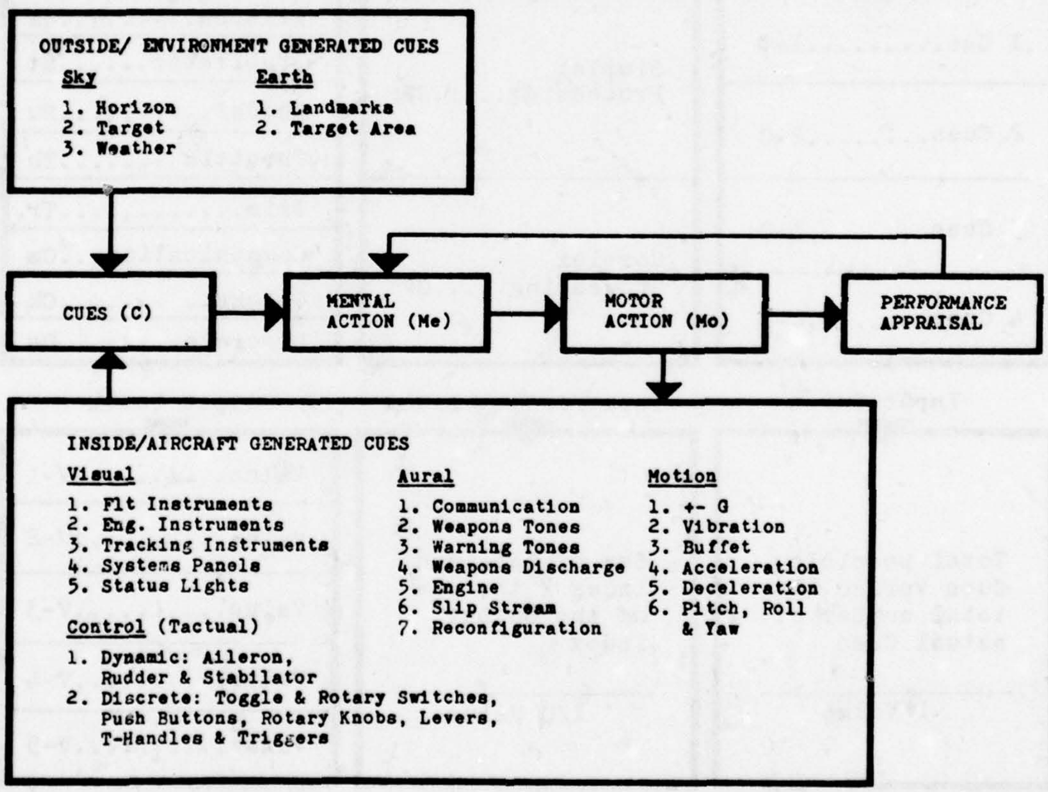


FIGURE 1

1 CUES

Kind

- Visual.......V
- Aural........A
- Control......C
- Motion.......M

Quantity

- 1 Cue.........1-C
- 2 Cues........2-C
- 3 Cues........3-C
- 4 Cues........4-C

Input Index

Total possible Cues versus the total number of actual Cues

I Value

2 MENTAL ACTION

Information Process

- Multi-Cue...........MC (Determines)
- Iterative...........I (Sustains)
- Specific-Cue........SC (Discerns)
- Multi-Cue/Iter......MC(I) (Determines/Sustains)
- Memory Recall/Iter..MR(I) (Anticipates/Sustains)
- Specific-Cue/Iter...SC(I) (Discerns/Sustains)

Decision Process

- Simple Processing.....SP
- Complex Processing.....CP

Input/Output Index

Sum of the input index X the sum of the output index

I/O Value

3 MOTOR ACTION

Continuity

- Establish Attitude......A
- Establish Rate of Attitude Change........R

Motor Output

- Aileron.........Ai
- Stabilator......St
- Rudder..........Ru
- Throttle........Th
- Trim............Tr
- Communication...Cm
- Checks..........Ck
- Discrete........Ds

Output Index

- Value..........V-1
- Value..........V-2
- Value..........V-3
- Value..........V-4
- Value..........V-5

FIGURE 2

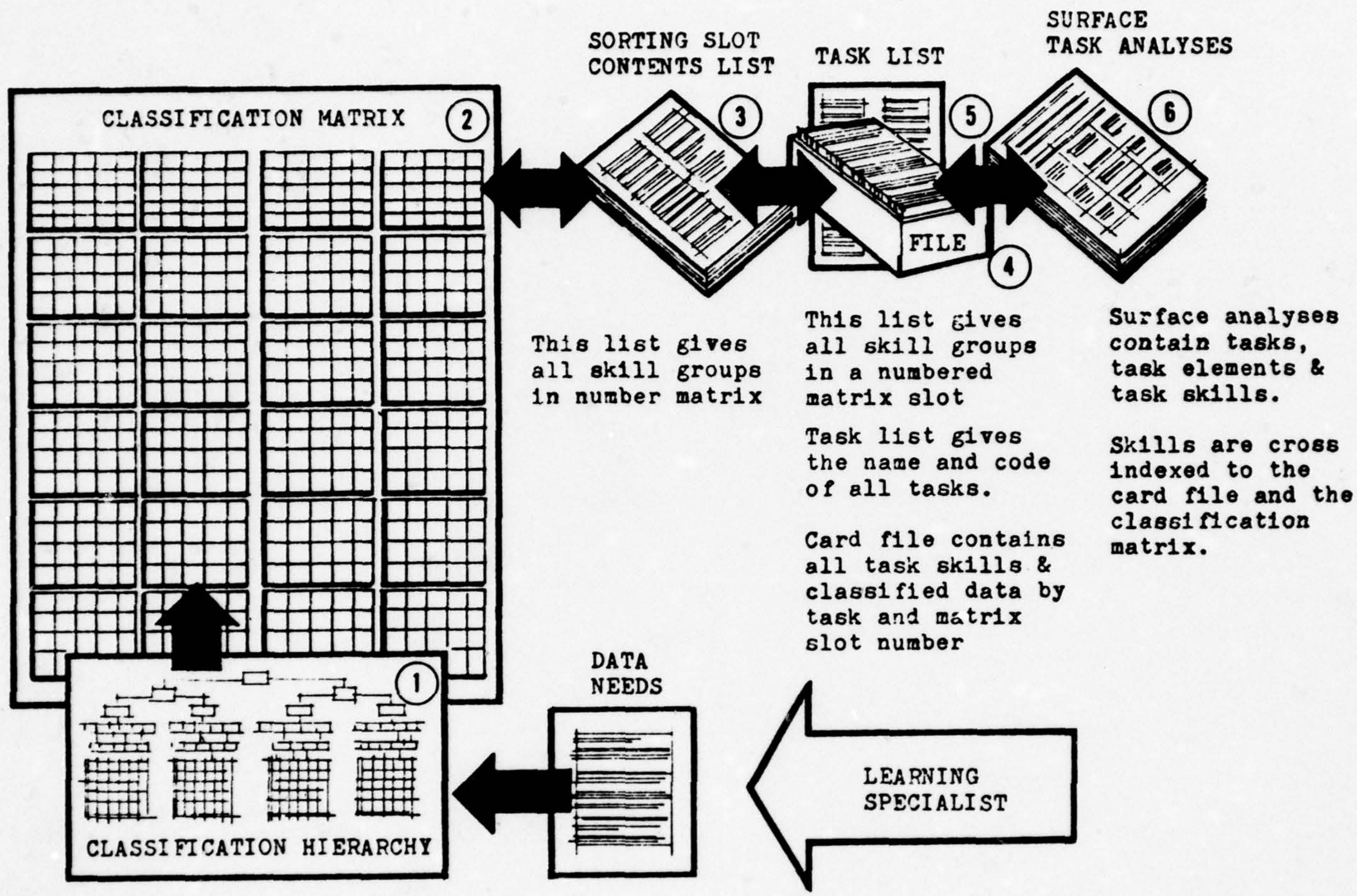

CLASSIFICATION MATRIX
2
SORTING SLOT
CONTENTS LIST
3
TASK LIST
5
FILE
4
SURFACE
TASK ANALYSES
6
This list gives all skill groups in number matrix
This list gives all skill groups in a numbered matrix slot
Task list gives the name and code of all tasks.
Card file contains all task skills & classified data by task and matrix slot number
Surface analyses contain tasks, task elements & task skills.
Skills are cross indexed to the card file and the classification matrix.
1
CLASSIFICATION HIERARCHY
DATA
NEEDS
LEARNING
SPECIALIST

FIGURE 3

VISUAL DISCRIMINATION PRETRAINING
OF A UPT LANDING MANEUVER

Bernell J. Edwards
Bruce A. Smith
and
Douglas C. Weyer
Human Resources Laboratory
Williams Air Force Base

PROBLEM

Undergraduate pilot training instructor pilots generally agree that the final turn of the T-37 overhead landing pattern is the most difficult training task in the early stages of UPT. Eventually, the experienced pilot will learn to fly this maneuver using visual references almost exclusive of the cockpit. IPs typically try to direct the student's attention to external references as soon as possible. As a training problem, the final turn may be considered a likely focus for improved training methods.

Normally, the student learns the visual aspects of this maneuver by acquiring proper responses to the discriminative stimuli within the visual field. Through this trial-error process, he becomes able to distinguish correct from incorrect altitude, ground track, runway orientation, and to associate appropriate control inputs to maintain the proper flight path. All of this is highly dependent upon proper use of the visual field. Thus, specific training in the visual aspects of the maneuver might be an aid to the development of judgment and skill required to fly the turn. An effective method of pretraining salient visual discriminations prior to cockpit time might increase pilot effectiveness and reduce required training time.

The basic features of visual discrimination theory as a learning method have been documented in the work of Bijou; Terrace, Lumsdaine, Sulzer and Kopstein; Ellis, Gagne and Baker; Goss and Greenfield; Norcross and Spiker; Vanderplas; and others. The evidence suggests that training of discriminative visual stimuli may substantially aid the acquisition of visual perceptual motor skills associated with those visual stimuli.

The objective of the present study was to develop and evaluate the pretraining of visual skills involved in the final turn landing pattern. Three phases were required: (1) development and validation of a visual discrimination pretraining (VDPT) strategy, (2) application of the strategy to determine transfer effectiveness for both the simulated and real time environments of T-37 training, and (3) evaluation of the strategy for broadest possible impact in flying training.

Extensive development of audiovisual media was undertaken to produce a self-paced instruction program. The program consisted of a series of wide-angle, rear projected color transparencies of the final turn as seen from the cockpit from a variety of positions throughout the turn, and from various deviations from the normal turn path. Pictures were sequenced so that the student viewing the program could compare simultaneously various aircraft positions by contrasting slides showing a deviant position with a slide showing a normal position at the same point along the turn path. Levels of difference throughout the program were arranged in order of difficulty. Textual materials, pointer arrows, and pictorial insets of cockpit instruments supplemented each of the visual comparison displays. The program sequence featured frequent review and practice so that *S* could assess his own progress in acquiring the visual skills. He controlled the pace of the instruction and was able to reverse the order of slides and review material at will.

A 63-item visual discrimination achievement test was constructed to assess the effectiveness of the program. The test required *S* to identify a series of slides of the final turn flight path according to various possible deviations from the correct path. The program and test were revised and validated using instructor pilots and naive casual students awaiting entrance into UPT at Williams AFB.

METHOD

Thirty-eight male student subjects from Class 77-08 at Williams AFB participated in an experiment to determine the effectiveness of the VDPT materials. Prior to the experiment, *Ss* had received no training in the overhead landing pattern. The random assignment of *Ss* to treatment groups was as follows: group one control, received no special training; group two received pretraining in the procedural aspects of the final turn and saw slides of the normal flight path, but received no VDPT; group three received the VDPT; group four received no pretraining but was subjected to final turn maneuver practice in the Advanced Simulator for Pilot Training (ASPT); and group five received both VDPT and ASPT. All groups received aircraft training in the final turn as a regular part of the UPT syllabus.

Versions of pretraining (procedures only for group two and VDPT for groups three and five) were administered just prior to the start of flight line training in the landing pattern. *Ss* received the two-hour self-paced pretraining package and returned the following day for the 30-minute achievement test. As was expected, both groups receiving the VDPT (three and five) scored significantly higher ($p < .01$) on the test than did the procedures only group two.

To assess the transfer effects of VDPT on execution of the final turn in a simulated environment, group four (ASPT only) and group five (VDPT plus ASPT) were compared. Each *S* received two one-hour sorties with an IP on two consecutive training days prior to flight line training in the overhead pattern. He was given a brief orientation to the simulator and a brief warmup period for familiarization. *S* was then given three no-prompt attempts to execute the maneuver to assess any direct transfer of VDPT. Next, *S* was shown a computerized model demonstration of the final turn with instructions from the IP. *S* then attempted repeated trials in flying the turn. During and between trials the IP provided prompts and feedback on performance. The type and frequency of prompting was controlled to preclude differences between *Ss* and groups. The total of final turn trials for each student was 24, 12 per sortie.

Dependent measures of *S* performance in ASPT were automated scoring of deviations (RMS error) from acceptable tolerance bands for altitude, bank angle, and airspeed. Subjective IP ratings were also used as a dependent measure for selected aspects of the maneuver.

The 24 trials flown by each *S* were grouped into three-trial groups making a total of eight data groupings for each student on each flight parameter: altitude, bank angle, and airspeed. RMS error mean scores for each *S* on each parameter were then analyzed using a two by eight repeated measures ANOVA (Winer, 1971).

Error comparisons for altitude and bank angle showed no significant differences. The altitude deviation error for the non-pretrained group (four) was observed to be significantly less than that of the pretrained group (five).

On the third trial of each three-trial grouping *S* was permitted to extend the flight path through the final approach segment of the pattern. RMS error data for ground path position, course, and airspeed were analyzed using a one-way ANOVA to determine any performance differences between the groups on these performance aspects. No statistically significant differences were observed between the two groups on these measures.

IPs rated performance of *Ss* as satisfactory or unsatisfactory on critical aspects of the final turn maneuver in ASPT: adherence to correct ground path, starting position for the turn (altitude and ground path combination), altitude and airspeed during the first and second halves of the turn. Dichotomous performance ratings for each of these elements were analyzed with chi square tests. No statistically significant differences between groups were observed for this data.

Transfer of VDPT to Aircraft Training

Flight line IPs rated their students' performance (all *S* in all groups) on selected critical elements of the final turn across the first ten turns attempted as a means of determining the effects of VDPT on performance in the aircraft. Five performance elements rated as satisfactory or unsatisfactory were as follows: (1) altitude at start of turn, (2) ground track at start of turn, (3) ground track during turn, (4) altitude during turn, and (5) airspeed during turn. IPs also gave a composite or overall rating for each final turn trial. The general rating was based upon a nine-point scale from unsatisfactory to excellent. Ratings for the five performance elements were analyzed by chi square tests; the general rating by a Kruschal-Wallis test (Siegal, 1956).

The chi square test for three of the five performance elements rated revealed statistically significant differences between groups. The among group difference for altitude at start of turn (chi square = 12.63) was significant at the .05 level with significantly better ratings, as revealed by post hoc chi squares for the following between group comparisons: group one (control) over group three (VDPT), $p < .005$; group four (ASPT) over group three (VDPT)), $p < .05$; and group five (VDPT/ASPT) over group three (VDPT)), $p < .01$.

The among groups ratings comparison for the ground track at start of turn performance element (chi square = 15.98) was significant at the .01 level with significantly better group performance found in the following sub-comparisons: group one (control) over group three (VDPT), $p < .01$; group two (procedures only) over group three (VDPT), $p < .001$; group four (ASPT) over group three (VDPT), $p < .05$; and group five (VDPT/ASPT) over group three (VDPT), $p < .005$.

The general comparison for ground track during final turn (chi square = 9.76) was significant at the .05 level with post hoc chi square tests showing significantly better performance for group one (control) over group two (procedures only), $p < .05$ and group one (control over group three (VDPT), $p < .05$.

The results of the Kruscnal-Wallis test for overall turn performance ratings revealed no reliable differences among groups.

To summarize in aircraft training results, there were no instances observed where the VDPT group was rated superior to other groups. This shows that the VDPT did not produce the expected training benefits. In fact, VDPT group performance in these comparisons was inferior giving evidence of a negative transfer effect of VDPT. Further, there were no data to show that group four (ASPT) or group five (VDPT/ASPT) performed elements of the turn better than did the control group. However, groups four and five were rated significantly better than the VDPT group on two performance elements. This may suggest that for these aspects of performance the ASPT training combined with VDPT may have aided in overcoming potentially deleterious effects of the VDPT.

DISCUSSION

Results of this study do not demonstrate a clear benefit for visual discrimination pretraining. Although inconsistent across all measures, there is evidence of a negative transfer of training effect. There are several possible reasons for this outcome.

Extensive elaboration of external visual cues at an early point in training may be inappropriate to student readiness. It is well known that the novice pilot tends to reference in-cockpit cues and may merely become confused by early emphasis of the external environment. In order to facilitate the integration of the external visual environment and aircraft control skills a different application of the VDPT might be effective. Rather than a single "dose", a distributed schedule of VDPT across some number of landing practice sorties should permit the integration of visual and motor skill components. In such a training scenario VDPT effects would appropriately be measured cumulatively across a substantial number of sorties rather than a "one time" effect early in the training. Also, the content of VDPT might be expanded to include a larger portion of the landing task such as the downwind and final approach in addition to the final turn as a means of integrating landing skill components temporarily.

The results of the present study, if regarded in a formative sense, are useful. Enough has been learned from the present data, and enough further trial of investigation VDPT concept. A follow-on study is planned to substantiate and extend what has been learned from this first VDPT experiment.

REFERENCES

Battig, W. F., Transfer from verbal pretraining to motor performance as a function of motor task complexity. *Journal of Experimental Psychology,* 1959, *51,* 371-378.

Canton, G. N., Effects of three types of pretraining on discrimination learning in preschool children. *Journal of Experimental Psychology,* 1955, *49,* 339-342.

Eddowes, E. E., McRee, P. L., Matheny, W. G., & Crowder, N. A., *Preliminary operational evaluation of an audiovisiual instrument trainer.* AFHRL-TR-75-49, AD-A016 487. Williams Air Force Base, AZ: Flying Training Division, Air Force Human Resources Laboratory, July 1975.

Ellis, H. C. and Muller, D. G., Transfer in perceptual learning following stimulus predifferentiation training. *Journal of Experimental Psychology,* 1964, *68,* 388-395.

Ellis, H. C., *The transfer of learning;* New York: The MacMillan Company, 1965. Gagne, R. M. and Baker, K. E., Stimulus Predifferentiation as a factor in transfer of training. *Journal of Experimental Psychology,* 1950, *40,* 439-451.

Goss, A. E. and Greenfield, N., Transfer to a motor skill as influenced by conditions and degree of prior discrimination training. *Journal of Experimental Psychology,* 1958, 3:55, 258-269.

Norcross, K. J. and Spiker, C. C., Effects of mediated association on transfer in paired-associate learning. *Journal of Experimental Psychology,* 1958, *55,* 129-134.

Siegal, S., *Non-parametric Statistics:* New York: McGraw-Hill Book Company, Inc., 1956.

Smith, B. A., Waters, B. K. and Edwards B. J., *Cognitive pretraining of the T-37 overhead traffic pattern.* AFHRL-TR-75-72, AD-A026 118. Williams AFB, AZ: Flying Training Division, Air Force Human Resources Laboratory. December, 1975.

Vanderplas, J. M., Sanderson, W. A. and Vanderplas, J. N., Some task-related determinents of transfer in perceptual learning. Perceptual Motor *Skill,* 1964, *18,* 71-80.

Winer, B. J., *Statistical Principles in Experimental Design;* New York: McGraw-Hill Book Company, Inc., 1971.

THE ROLE OF THE BEHAVIORAL SCIENTIST IN APPLIED RESEARCH

Ronald L. Johnson and Edward R. Eichelman
U.S. Army Mobility Equipment Research and Development Command

This study demonstrates the role of the behavioral scientist in controlled quantitative research. Thirty-five two-man teams of operational image interpreters analyzed visual imagery containing standard and toned down 155mm ammunition packaging material. A chi-square and co-efficient of association analyses performed upon probability of detection indicated that tone down colors were significant at the .05 level in reducing pallet detection. The treatment was not significant for shipping containers on pallets. The apparent conflict of the effects of tone down was solved by the behavioralists. This was accomplished through the subjective evaluation of visual cues reported by the operational image interpreters as aiding in target detection. The addition of texturing material to the tone down colors was recommended to disperse reflected light from curved surfaces.

The US Army Training and Doctrine Command, in a letter [1] dated 2 Dec 1975, stated that the detection of battlefield litter was a potential aid to enemy observers and sensors in the acquisition of targets. This letter along with a feasibility study [2] of tone down signature reduction formed the basis for this paper. Behavioral scientists were given the task to examine the following test hypothesis:

The use of tone down colors will reduce the detectability of 155 mm ammunition packaging material.

Method

Past evaluations of camouflage have shown that operational image interpreters produce the most reliable and reproducible quantitative data. This study was conducted in two phases. The required aerial imagery was taken in the first or field phase. The imagery was given to image interpreters for analysis in the second or laboratory phase.

Field

A Fort Belvoir, VA, training area, T-9, was selected as the test site. The site supported a vegetation cover of widely spaced pine and deciduous hardwood trees varying in height up to 6 m. Ground cover consisted of short brush and grass varying in color from light to dark green. The bare earth was a dark reddish brown. Imagery was obtained with a Zeiss RMK-15-23 camera system, employing a 6 inch focal length lens. The camera system was mounted in a fixed wing Aero Commander. Nine inch format, Kodak color Ektographic film number 2448 was taken at a scale of 1:5,000, covering a land area of 124 acres per frame. Each frame was taken in 60 percent forward overlap to enable the use of stereoscopics. The pallets and the shipping containers on pallets were placed in the test site with no attempt to physically change or mask their characteristic geometric shapes. They were first photographed in the untoned condition and then with color tone down. The resulting imagery was cut into two 12 frame strips of film precluding center and end target location.

Laboratory

A total of 35 two-man teams of military, operational image interpreters stationed along the East coast participated in the study. Eighteen teams viewed one strip each containing the standard pallets and shipping containers on pallets. The remaining 17 teams viewed the second strip of imagery containing tone down pallets and shipping containers on pallets. Each team was given one hour to analyze the imagery. They were instructed to search for packaging material. As the interpreters detected each potential target, they circled it with a grease pencil and numbered it in order of detection. They immediately reported all visual detection cues to the behavioral scientist who used the free response or open-ended interview technique. The probability of detections were also determined and recorded.

Results

A chi-square analysis [3] and a co-efficient of association [4] were performed upon the detection data. These results are seen in tables 1 and 2.

Discussion

From a statistical aspect, the test hypothesis was acceptable according to chi-square values of Table 1, and not acceptable according to the values of Table 2. Thus, upon first appearances, one may be tempted to conclude that nothing was proven by the study, and add it to the numerous volumes of non-significant results. However, behavioral scientists' analysis of the data (Table 3) indicates that the use of tone down colors was indeed significant in reducing the probability of target detection. The primary visual cues leading to target detection of the standard pallets and shipping containers on pallets were geometric shape and lighter color than the background. Both of these cues virtually disappeared when tone down was added to the standard pallets. The response of lighter color had the same frequency for the standard pallets, standard shipping containers on pallets, and tone down shipping containers on pallets. Since identical pallets were used with and without shipping containers, it was hypothesized that the problem of detections must rest with the shipping containers. This hypothesis was also supported by the frequency of the response, geometric shape, for the toned and untoned shipping containers on pallets. The curved surfaces of the shipping containers reflected light back to the camera system. This reflection was reported by the image interpreters as a light color. The addition of a texturing material to the tone down colors would diffuse the reflected light. Thus, the scientific techniques employed by the behavioral scientist has enabled a positive conclusion. The use of tone down plus the addition of texturing will reduce the probability of target detection.

TABLE 1
Statistical Summary of Detections of Standard & Tone Down Pallets

Test Items	Detected	Not Detected	Total
Standard Pallets	10	8	18
Toned Down Pallets	2	15	17
Total	12	23	35

$\chi^2 = 5.54$ $\chi^2_{.95(1)} = 3.84$

Reject Null Hypothesis of Equal Proportions.

Co-efficient of association (Q) = .807.

Accept test Hypothesis that the use of tone down colors reduced the detectability of the pallets.

TABLE 2
Statistical Summary of Detections of Standard and Tone Down Shipping Containers on Pallets

Test Items	Detected	Not Detected	Total
Standard Shipping Containers on Pallets	7	11	18
Tone Down Shipping Containers on Pallets	9	8	17
Total	16	19	35

$\chi^2 = 0.24$ $\chi^2_{.95(1)} = 3.84$

Cannot reject Null Hypothesis of Equal Proportions.

Co-efficient of association (Q) = -.277, N. S.

The frequency of visual cues reported by the image interpreters as aiding in terget detection of all targets are recorded in table 3.

TABLE 3
Frequency of Use of Visual Cues In Target Detection

Visual Cues	Test Targets			
	A	B	C	D
Geometric Shape	9	1	4	5
Lighter Color	7	1	7	7
Shadows	–	–	1	4
No Shadows	1	1	–	–
Large Size	–	1	1	–
Height	–	–	2	4
No Height	2	1	–	–

KEY A - Standard Pallets
B - Tone Down Pallets
C - Standard Shipping Containers on Pallets
D - Tone Down Shipping Containers on Pallets

REFERENCES

1. Letter, ATCD-SP-L, TRADOC, Subject: *Camouflage Requirements for Packaging Materials*, 2 Dec 75.
2. Hansen, Clay M., and Spilman, James M., *Feasibility Study of Camouflage Material, 155 mm Ammunition*, Picatinny Arsenal, Dover, NJ.
3. Dixon, Wilfred, J., and Massey, Frank J., Jr., *Introduction to Statistical Analysis*, 3d Edition, McGraw Hill, 1969.
4. Moroney, NJ, *Facts From Figures*, Pelican Book Co., 1952.

OPERATOR WORKLOAD ASSESSMENT IN SYSTEM TEST AND EVALUATION

Wade R. Helm
Pacific Missile Test Center

Adequate assessment of operator workload in a complex system requires a substantial testing program which in turn yields large amounts of data. The different pieces of data regarding operator workload have different relative importance in terms of implications for required decisions such as system redesign or compensatory training. Since no single metric exists that summarizes the data along the relevant dimension, a means must be employed to gather and organize data in a comprehensive fashion that facilitates decision making about such factors. This paper reports on the feasibility of the application of computer aided evaluative techniques in combining precise pieces of test information into summary measures at differing levels of generality to facilitate decisions about operator workload and system acceptability.

The adequate assessment of operator workload in highly complex aviation systems requires the collection, reduction and evaluation of extremely large amounts of data. Collecting insufficient data results in an assessment of workload that is too general in nature to provide the detail necessary to make decisions concerning crewstation redesign or the application of appropriate training techniques. Collecting enough data to answer the redesign or training questions requires that the data be organized and structured in such a manner that provides the decision maker the detailed task level assessments in relation to higher level responses encompassing operator behavior. If the data is not integrated from simple tasks to complex levels of behavior, there will be no means available to adequately assess the relative contributions of each task to overall operator workload.

To obtain workload information that will reasonably reflect task difficulty during a mission scenario, the data must be collected in operational or simulated operational environments. Currently one of the few available assessment methodologies comprehensive enough to cover the entire spectrum of operator activities is task or job analysis. Task or job analysis can provide the necessary framework or structure of operator activities during an entire mission profile. Combined with task analysis, scaling methodology can provide the means of combining inputs from different operators as to task difficulty and subsequent workload assessment. This approach not only provides data on tasks that operators have determined are difficult to accomplish but also provides reliability estimates of operator ratings. However, this approach is limited in that there is no way to evaluate comprehensively the impact that each task has on total operator workload.

One solution to the problem of integrating tasks into a comprehensive workload index requires the adaptation of multi-attribute utility theory methodology. Employing a hierarchical multi-attribute utility model combined with an interactive computer system provides a means of assessing the dynamics of task requirements in a complex system. For example, this methodology takes related pieces of specific data and organizes the information into a hierarchical structure so that each successive higher level in the hierarchy there is a greater degree of generality. Decision makers can then work at that particular level of the structure that is best suited to the nature of the decision required. The approach has two unique elements. One, data obtained at the task level can be integrated according to rules relative to multi-attribute utility theory, and, two, the data at all levels of the hierarchy are made readily accessible through a computer interactive system.

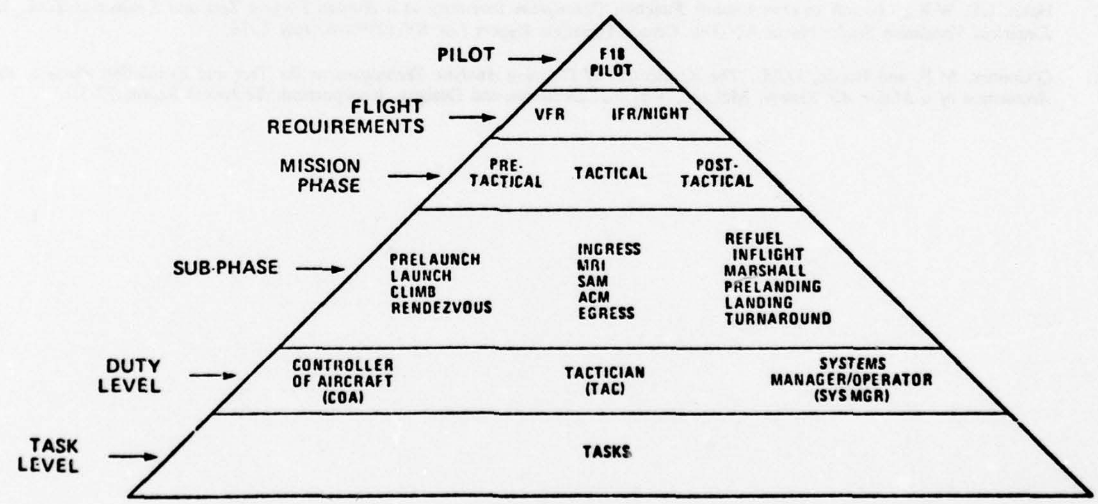


FIGURE 1
THE F18 PILOT TASK INVENTORY HIERARCHY

A current hierarchical structure under devlopment for the Navy's F-18 fighter system is displayed in Figure 1. As illustrated, the structure contains six levels. The top level is the summary level for the entire aircraft, the pilot level. The next three levels contain flight requirements, mission phases, and mission sub-phases which are well documented in a scenario prepared to accompany the test flight evaluation The next level, the duty level, described the roles the pilot must play, i.e., controller of aircraft, tactician, and manager of the aircraft systems. The lowest level contains specific tasks that must be accomplished dependent on mission subphase and duty.

In conjunction with hierarchical structure development, a concurrent effort is the development of an appropriate rating scale(s) to rate the various tasks and mission phases. Such a rating scale must have numerous qualities: (a) It must promote inter-rater consistency; (b) It must promote inter-rater agreement; (c) The scale must be valid. But above all the scale categories must be clear, meaningful, and directly relatable to observed aircraft properties.

The problems inherent in scale development are compounded by the complex nature of the pilot's job. For example, the tasks a pilot must accomplish are actually multi-dimensional and involve many sub-tasks. The difficulties in accomplishing tasks can therefore vary due to the nature of the aircraft and the skills required. One task may require physical efforts beyond that of most pilots, another may involve a tremendous cognitive workload, one that exceeds the capabilities of skilled test pilots, let alone those of less experienced pilots. Any valid measure of workload must reflect this interdependence between task difficulty and system design effectiveness. A measurement scale that is being developed to measure these two dimensions is presented in Figure 2.

The development of this scale is not yet complete but when implemented they will allow pilots to rate F-18 tasks on two simple, apparently ordinal scales which, based upon previous knowledge, will be combined to form a complex interval scale. This scale output will refelct pilot workload for that task rated.

Summary

Preliminary investigations on the F-18 indicates that it is possible to combine task analytic techniques and a multi-attribute utility model to develop a comprehensive workload assessment of a complex mission system. An F-18 task analysis and evaluative computer program has been completed. Following data analysis, an improved F-18 task analysis and a computer evaluative program will be constructed and available by Fall 1979. Using this approach, data will be collected during F-18 flight tests and the results will be presented to the F-18 program manager through an interactive computer system. With this system the program manager can determine overall workload assessments and, more importantly, how specific operator activities have contributed to total workload. With this information the program manager can determine which changes to the F-18 will be most beneficial in reducing operator workload.

REFERENCES

1. Cooper, George E. and Harper, Robert P., *The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities*. National Aeronautics and Space Administration, NASA TN D-5153, April 1969.

2. Helm, LT. W.R., *Third Interim Report: Human Factors Evaluation of Model P-3C Update 1 Airplane*. Naval Air Test Center Technical Report No. SY-122R-75, February 1975.

3. Helm, LT. W.R., *Fourth Interim Report: Function Description Inventory as a Human Factors Test and Evaluation Tool: An Empirical Validation Study*. Naval Air Test Center Technical Report No. SY-127R-76, July 1976.

4. O'Connor, M.F. and Buede, D.M., *The Application of Decision Analytic Techniques to the Test and Evaluation Phase of the Acquisition of a Major Air System*. McLean, Virginia: Decisions and Designs, Incorporated, Technical Report 77-3.

TECHNICAL EFFECTIVENESS	H	G	F	E	Column / Row
Design enhances mission phase accomplishment, multiple tasks integrated					A
Design enhances specific task accomplishment					B
Adequate performance achievable					C
Inadequate performance due to technical design, cannot compensate for sub-par performance					D
	Workload at critical level; Interference with A/C control; Compensation very excessive; Dangerous	Workload considerably higher than anticipated; Compensation high; Interference major	Workload slightly higher than anticipated; Moderate compensation; Minor interference	Workload as anticipated; No Interference; No compensation	

PILOT WORKLOAD

FIGURE 2

BIORHYTHMS AND THEIR RELATION TO PILOT ERROR ACCIDENTS AND INCIDENTS

Hal W. Hendrick
Harlan E. Jones
University of Southern California

In spite of the tremendous advances in the safety technology of aircraft design and manufacture, development and application of human factors engineering, sophistication of pilot and crew training simulators and methodology, and rigorous safety programs, aircraft accidents and incidents attributable to pilot error continue to persist with nagging regularity. To date, identification of underlying causal factors of these human errors has remained elusive. The often heard superficial explanations, "such as the pilot just wasn't up to speed", or that "he was behind the power curve" simply do not suffice, but they do suggest a potentially fruitful, though controversial direction for investigation, the study of biorhythms.

The pioneers of biorythm research were Dr's Herman Swoboda of Vienna and Wilhelm Fliess of Berlin. Their research at the end of the 19th century consisted of studying the periodicity of pain and tissue swelling in their patients. Both men wrote extensively on their findings that nature had seemingly given to man at birth a biological clock. They concluded that there was a 23 day rhythm which affected persons physically, and a 28 day rhythm which affected the emotions. Later, in the 1920's Professor Alfred Teltscher at the University of Innsbruck discovered a 33 day intellectual cycle by studying the performance records on exams of hundreds of high-school and college students. (Halberg, 1975)

These three biorhythms begin their continuous cycles from the time of birth and continue until death. From birth the cycles begin a positive rise from a neutral or zero point. They eventually reach an apex, and then decline through the zero point until they reach a trough or low point; from here they rise and ascend back through the zero point, completing the cycle. When a given cycle crosses the zero point, it is referred to as a *critical day*. Contrary to popular belief, an individual's weak or unstable days occur not during the trough or negative phase, but on these critical days when the individual is in a period of phase transition (Gittelson, 1976). The 23 day physical rhythm was believed by Fliess to originate in the muscular cells; fluctuations of the rhythm appeared to affect one's strength, endurance, and resistence to infection. The 28 day emotional cycle appears to be associated with the rise and fall of estrogen, progesterone, and aldosterone; during their positive phase persons have been found to be relatively positive and cheerful, whereas during their negative phase they tend to be less optimistic and more irritable. The 33 day cycle appears to affect learning and analytical processes. Gittleson (1976) has noted that while there is little scientific proof that these rhythms exist in some tangible form, there are literally hundreds of studies which document the occurance of these 23, 28 and 33 day rhythms in cyclic form.

The major application of biorhythm theory has been in the field of accident prevention and investigation. In Japan, where a grop of physicians, insurance executives, and safety engineers have founded the influential Japan Biorhythm Association, utilization of biorhythm studies in safety programs is widespread. Over 5000 firms are now using this technique, including the well known Hitachi Manufacturing, Fuji Heavy Industries, Mitsubshi Heavy Industry, Bridgestone Tire, Japan Express, and Japan Air Lines. These companies have instituted biorhythm safety programs which reportedly have resulted in accident reductions of 35 to 45 percent. In these programs, each person receives or makes his own biorythm chart and on a critical day will receive a warning slip advising extra caution on that day. A similar program was introduced by Mike Bertalot of United Airlines in 1973, and now includes between 6000 and 8000 United employees. (Cavenough, 1975)

In the late 1930's Dr. Hans Schwig, a behavioral scientist at the Swiss Institute of Technology, conducted the classic study relating critical biorhythm days to accident injuries and deaths. From review of 1400 accident case histories obtained from Zurich insurance companies, Schwig found 40 percent of the injuries and 65 percent of the deaths due to accidents occurred on critical biorhythm days (Nelson, 1976). Nelson notes that similar results were recently reported by Jacob Sanhein of the Naval Weapons Support Center during the fall 1975 Congress of the National Safety Council. In his study, Sanhein found over 40 percent of the accidents to have taken place during the 20.4 percent of the time comprised of the victim's accident prone days.

The study of the relation of biorhythms to aircraft accidents where human error was suspected also has received some recent attention. In particular the Guggenheim Aviation Safety Center has found nearly 80 percent of the private avaiation suspected pilot error accidents reviewed to have occurred on a critical biorhythm day for the pilot (Nelson, 1976). Gittleson (1976) reports that in 10 of 13 recent major plane crashes attributed to pilot error, the pilot or co-pilot was on a critical biorhythm day.

In the present study, it was hypothesized that aircraft accidents and incidents attributed to pilot error would be found related to critical days for one or more biorhythms of the pilots involved.

METHOD

Subjects.

The subjects in this study consisted of two groups of 25 male military pilots who had been involved in an aircraft accident and one group of 50 male military pilots who had been involved in an aircraft incident which was attributed to pilot error based on routine safety investigations. These pilots ranged in rank from Lieutenant through Colonel and in age from 24 to 46 years. Their pilot experience varied from 500 to over 5000 hours of flight time. All 100 pilots were from the same branch of military service.

Apparatus.

Biorhythm phase for each pilot at the time of his accident or incident was calculated using the Casio "Biolator" model H-801. This is an electric calculator designed expressly for the purpose of determining to the nearest day one's biorhythm phase for all three biorhythms. The Casio H-801 instruction manual defines the day to either side of the zero day as semi-critical days, and they can be considered along with the zero day to comprise the critical phase. Each biorhythm cycle has two critical phases of three days each.

Procedure.

The accident investigation reports for a large military aviation unit which had experienced a dramatic increase in pilot error accidents during the 1976 and first half of 1977 period were reviewed, and the pilot's birthdate and accident date were extracted for each aircraft accident attributed to pilot error. For this 18 month period, 25 pilot error accidents were identified. A cross validation sample was selected by sequentially reviewing the accident reports for 1975 and earlier until 25 pilot error accidents were identified; these occurred during a 42 month period from July 1972 through 1975. A second cross validation sample consisting of 50 pilots who had been involved in aircraft incidents attributed to pilot error was identified by reviewing incident reports for the same military organization from November 1976 through May 1977.

For each pilot in all three groups, the biorhythm phase at the time of accident or incident was calculated for all three biorhythms. For each group, a Chi Square test for independence was conducted for the frequency of occurance of critical and noncritical phases for all three biorhythms and for the occurance of double critical phase biorhythms. Yate's correction was used in all cases where the expected frequency for one or more cells was less than 10.

RESULTS

The results for the validation accident group are summarized in Table 1. In only one case did the observed frequency of critical days exceed chance at or beyond the .05 level, and that was for the physiological biorhythm where the Chi Square was significant beyond the .01 level. Whereas 6.5 of the accidents could be expected to occur on a critical physiological biorhythm day by chance alone, twice that number or 13 of 25 actually occurred during this phase.

Table 2 summarizes the results for the cross validation accident group. Again, only the Chi Square for the physiological biorhythm reached significance at or beyond the .05 level. For this group, 12 of the 25 accidents occurred during a critical physiological biorhythm for the pilot.

The results for a second cross validation group are summarized in Table 3. This group was composed of 50 pilots who had been involved in an aircraft incident rather than an accident. As with the other two groups, only the Chi Square for physiological biorhythm reached significance. Here, 20 of the 50 incidents occurred during the pilot's critical physiological phase as compared to 13 which could be expected by chance.

Summarizing the physiological biorhythm data across all three samples, 45 of the 100 accidents and incidents occurred during the pilot's critical phase, whereas only 26 would have been expected by chance. These data yield a Chi Square of 18.8, which is significant beyond the .001 level. For this entire group the Chi Square tests of the frequency of occurrance of both positive and negative biorhythm phases also failed to reach significance for all three biorhythms.

DISCUSSION

The significant finding for physiological biorhythm is consistent with the results of an organizational assessment survey of all pilots within this same military organization conducted in late 1976. In the survey results, pilots consistently reported feeling pressure to fly after a full day of office work or at other times when they did not feel up to it. Follow-up interviews with a dozen pilots included in this biorhythm study consistently confirmed that on the day of their accident or incident, they had felt either physically or psychologically not up to flying, but had flown anyway because of perceived command pressure.

The consistent relationship of physiological biorhythm critical phase to aircraft accidents and incidents suggests this may be a fruitful area for more extensive investigation. Assuming follow-on studies yield similar results, a safety program similar to those used widely in Japanese industrial firms, cited earlier, might be an effective approach for reducing aircraft accidents and incidents.

TABLE 1
Summary of Data for Accident Validation Group

Biorhythm	Expected f	Observed f	χ^2
Physical			
Critical	6.5	13	
Noncrit.	18.5	12	8.7**
Emotional			
Critical	5.4	6	
Noncrit.	19.6	19	0
Intellectual			
Critical	4.5	4	
Noncrit.	20.5	21	0
Double			
Critical	4.7	7	
Noncrit.	20.3	18	0.9

**sig. at .01 level

TABLE 2
Summary of Data for Accident Cross Validation Group

Biorhythm	Expected f	Observed f	χ^2
Physical			
Critical	6.5	12	
Noncrit.	18.5	13	5.2*
Emotional			
Critical	5.4	6	
Noncrit.	19.6	19	0
Intellectual			
Critical	4.5	6	
Noncrit.	20.5	19	0.2
Double			
Critical	4.7	5	
Noncrit.	20.3	20	0

*sig. at .025 level

TABLE 3
Summary of Data for Incident Group

Biorhythm	Expected f	Observed f	χ^2
Physical			
Critical	13	20	
Noncrit.	37	30	5.1*
Emotional			
Critical	10.8	10	
Noncrit.	39.2	40	0.1
Intellectual			
Critical	9	8	
Noncrit	41	42	0
Double			
Critical	9.4	5	
Noncrit.	40.6	45	2.0

*sig. at .025 level

REFERENCES

Cavanaugh, Tom. Biorhythms in the sky. *Science Digest*, June 1975, pp 16-17.

Gittleson, Barnard. *Biorhythm -- A Personal Science*. Arco Publishing Co., Inc., New York, 1976.

Halberg, Franz. *Biological Rhythms* New York, Plenum Press, 1975.

Nelson, E. New facts on biorhythms. *Science Digest*, May 1976, pp 70-75.

SPEED AND ACCURACY IN LISTENING TO COMPRESSED SPEECH AS A FUNCTION OF PRIOR EXPERIENCE AND AUDING STRATEGY

Paul A. Gade
US Army Research Institute
Alexandria, Virginia

The present experiment attempted to assess the effect of different information seeking strategies on the rate at which individuals choose to listen to passages of time compressed speech, and on their comprehension of those passages. The effects on listening rates and on comprehension due to prior experience with compressed speech were also explored using 48 male and female Army enlisted personnel as subjects. No significant differences in speed or accuracy of performance were found among the various information seeking strategies employed. Preference data indicated that personnel prefer to listen to speech at rates well above normal speaking rates ($p<0.001$). Although prior experience with compressed speech did not influence preferred listening rates, it did influence the participants' listening rates when they were induced to listen to speech as rapidly as possible ($p<0.001$). Results are discussed in terms of Berlyne's epistemic curiosity hypothesis and in relation to the results of preference research by Lass, Foulke, Nester and Comerci (1976).

Speech compression devices enable audio recordings to be played back at rates faster or slower than originally recorded, without changing the pitch. The fact that untrained listeners can comprehend compressed speech has been well established in the literature. For example, Foulke (1968) showed that college students' comprehension scores did not decline significantly until word rate was increased above 250 word per minue (WPM). Shields (1975) demonstrated that Army Communication Processors could accurately identify the subject matter of highly technical communications compressed to 1.5 times their normal rate. Recently, deHaan (1977) found good comprehension in subjects who listened to historical material compressed to twice the normal speed.

Numerous studies have shown that subjects can be trained to listen to compressed speech at more than twice the normal rate, with no appreciable loss in comprehension (e.g., Grumpelt & Rubin, 1972; and Lambert, Shields, Gade & Dressel, in press). However, relatively little is known about how listening rate preferences are influenced by exposure to compressed speech or how listener rate preferences influence performance when listening to compressed speech. If the military is to make effective use of the technology of rate controlled speech, variables such as *information seeking strategies* and *listening rate preference* must be explored more fully.

In the present experiment, I was concerned with the influence of several variables on Army subjects' performance in self-paced learning situations using compressed speech. First, I wanted to know if instructions which fostered epistemic curiosity would produce the positive influence on memory for compressed speech that it has been shown to produce in reading (see Berlyne, 1954a, 1954b & 1960; Watts and Anderson, 1971; and Frick and Cofer, 1972, for a more detailed discussion of epistemic curiosity). Furthermore, would epistemic curiosity affect the rate at which subjects elected to listen to compressed speech? Finally, I wanted to explore the relationship between preferred listening rates and the maximum rates at which subjects could listen to compressed speech (see Lass, Foulke, Nester & Comerci, 1976; and Levine, 1975, for discussions of speech rate preferences).

METHOD

Participants

Participants were 37 male and 11 female Army enlisted personnel who were assigned to the Intelligence Command at Arlington Hall station. All participants had scores of 100 or better on the GT scale of the AFQT.

Apparatus

Participants were seated in an IAC accoustical chamber during the experiment. A Crown 800 variable speed tape recorder/player was used in conjunction with an AmBiChron (Koch, 1974) speech compressor/expander to present passages of speech. Participants listened binaurally to the speech through a pair of headphones. Participants controlled speech rates by manipulating a knob on a control box in the audio chamber. All participant responses were automatically recorded on a strip chart and on a small laboratory computer. This computer, an ADDS 1800-E, was used to regulate the onset and offset of the tape recorder, and to record listening times.

Stimulus Materials

Five passages of speech were selected from a Library of Congress recording (talking book) of *The Proud Tower* (Tuchman, 1966). A professional female reader read the passages at an average rate of 130 WPM. One of the passages was used as a practice passage. The remaining four passages were manipulated in pairs (AB and CD). Each pair was considered a single passage with respect to order effects and data analysis.

Design and Procedure

Participants were randomly assigned to one of the 16 cells of a 2^4 completely randomized factorial design. The four factors defining the cells were: 1) type of instructions (declarative or interrogative); 2) type of practice exposure to speech (normal or twice normal rate); 3) passage order (AB followed by CD, or the reverse); and 4) order of passage instructions (AB had instructions, while CD had no instructions, or the reverse). Upon arriving at the laboratory, participants were given a brief description of the experiment and the tasks they would be performing. They were then given a brief practice session during which they were shown how to control the rate of compressed speech. Next, the participants were told to listen to a practice passage, and that afterwards they would be asked questions about the passage.

They were also told that they would *not* be able to control the speed of the speech during this practice passage (and that they would have to listen to the passage at either normal speed or twice normal speed—depending upon group assignments). After they finished the questions at the end of the practice passage, participants were informed that they would next listen to four more passages for which they could control the rate of speech. They were asked to listen to each passage as rapidly as possible, but not so fast that they would be unable to answer the questions at the end of each passage.

Overall listening rates for each passage and the number or correct answers on each test served as response measures. After completing the test for the fourth passage, participants were asked to listen to that fourth passage again. While listening, they were asked to adjust the speed of the speech to the rate that they most preferred.

RESULTS AND DISCUSSION

In examining the effects of the information seeking strategy employed, two response measures were used: 1) number of correct answers in the tests following each passage; and 2) speed at which the participants listened to the various passages. A five-way ANOVA was performed on the correct answer data. It was expected that, based on the epistemic curiosity hypothesis, the comprehension scores for passages preceded by questions would be higher than those not preceded by questions. However, such was not the case; none of the main effects or interactions of primary concern were significant. The only significant effect was practice passage speed x presentation order x order of instruction x passage instructed interaction ($F(1,32)=5.58$, $p<0.025$). Subsequent simple effects testing using Tukey's *HSD* test (Kirk, 1968) failed to reveal any systematic differences between or among any of the groups of concern. In brief, in the present experiment, the instructions had no significant effect on comprehension performance.

A five-way ANOVA was performed on the speed at which subjects elected to listen to speech. A significant main effect for practice passage speed was found ($F(1,32)=20.01$, $p<0.001$). Subjects who were required to listen to the practice passage at twice the normal rate then elected to listen to the four subsequent passages at faster rates than those who listened to the practice passage at a normal rate ($\bar{X}=1.69$ times normal vs. $\bar{X}=1.39$ times normal, respectively). Two interactions were also significant: speed x order x passage ($F(1,32)=4.34$, $p<0.05$); and presentation order x order of instruction x passage ($F(1,32)=5.88$, $p<0.05$). However, as with the comprehension data, simple effects analyses for both these interactions failed to yield any meaningful differences. No other main effects of interactions were significant. A four-way ANOVA was performed on the preference data. No significant main effects or interactions were found; clearly indicating that the variables manipulated failed to affect participant preferences in any significant way.

It is interesting to note that 31 of 33 subjects preferred to listen to speech at a faster than normal rate ($X^2(1)=25.48$, $p<0.001$). Furthermore, the mean preferred rate (1.42 times normal rate) was significantly higher than the normal rate ($t(32)=10.24$, $p<0.001$). A two way ANOVA was used to analyze the differences between preferred and selected speeds for groups receiving the practice passage at twice the normal rate selected a significantly ($F(1,62)=8.92$, $p<0.05$) higher rate ($\bar{X}=1.66$ times normal) than the group receiving practice at the normal rate ($\bar{X}=1.43$ times normal) rate. However, these two groups did not differ in their preferred speech rates as the respective mean differences show (1.44 compared to 1.41 times the normal rate).

Furthermore, the group given practice at twice normal speed selected speeds ($\bar{X}=1.66$ *times normal) significantly higher* ($F(1,62)=13.90$, $p<0.001$) than their preferred speed ($\bar{X}=1.44$ times normal). No such differences between selected and preferred speeds were found for the normal speed practice group. Taken together these results seem to indicate that prior exposure to compressed speech modified the rate at which subjects elected to listen to further passages of speech without influencing their listening preferences in a significant way. Apparently selected and preferred listening rates were not entirely independent of one another since the correlation between selected and preferred rated was significantly different from zero ($r(31)=+.47$, $p<0.01$).

REFERENCES

Berlyne, D.E. A theory of human curiosity. *British Journal of Psychology*, 1954, *45*, 180-191. (a)

Berlyne, D.E. An experimental study of human curiosity. *British Journal of Psychology*, 1954, *45*, 256-265 (b).

Berlyne, D.E. *Curiosity, conflict and arousal.* New York: McGraw-Hill, 1960.

deHaan, H.J. A speech-rate intelligibility threshold for speeded and time-compressed connected speech. *Perception and Psychophysics*, 1977, *22* (4), 366-372.

Foulke, E. Listening comprehension as a function of word rate. *Journal of Communication*, 1968, *18*, 198-206.

Frick, J.W. & Cofer, C.N. Berlyne's demonstration of epistemic curiosity: An experimental re-evaluation. *British Journal of Psychology*, 1972, *63*, 221-228.

Grumpelt, H.R. & Rubin, E. Speed listening skill by the blind as a function of training. *The Journal of Educational Research*, 1972, *65*, 467-471.

Kirk, R.J. *Experimental design: Procedures for the behavioral sciences*. Belmont, California: Brooks/Cole, 1968.

Koch, R. The AmBiChron, In S. Duker (Ed.), *Time compressed speech: An anthology and bibliography*. Metuchen, New Jersey: Scarecrow Press, 1974.

Lambert, J.V., Shields, J.L., Gade, P.A. & Dressel, J.D. Comprehension of time-compressed speech as a function of training. Technical Paper, U.S. Army Research Institute for the Behavioral and Social Sciences, in press.

Lass, N.J., Foulke, E., Nester, A.A. & Comerci, J. The effect of exposure to time-compressed speech on subjects' listening rate preferences and listening comprehension skills. *The Journal of Auditory Research,* 1976, *5,* 179-186.

Levine, S.J. Preferences for the rate of presentation of recorded information. *Proceedings Third Annual Conference on Rate-Controlled Speech.* New York: American Foundation for the Blind, 1975.

Tuchman, B.,*The proud tower.* New York: Macmillan, 1966.

Watts, G.H. & Anderson, R.C. Effects of three types of inserted questions on learning from prose. *Journal of Educational Psychology,* 1971, *62,* 387-94.

BLOOD PRESSURE AND BEHAVIOR OF MALE AND FEMALE RATS ARE DIFFERENTIALLY ALTERED IMMEDIATELY AFTER EXPOSURE TO IONIZING RADIATION

G. Andrew Mickley
Armed Forces Radiobiology Research Institute
Bethesda, Maryland 20014

Abstract

Sufficiently large doses of ionizing radiation produce an early transient incapacitation (ETI) characterized by akinesia and a decrement in motivated behaviors. The etiology of this disorder is unclear but acute hypotension often accompanies the behavioral symptoms. Since administration of estrogens has been shown to prolong the survival of irradiated mice, behavioral and physiological differences between the sexes during this acute ETI were investigated. Females out performed males on an active avoidance task. In this same time period, males exhibited an acute hypotension while female blood pressure increased. Female rats may be more radioresistant than males.

Male rats which have received a sufficiently large and rapid dose of ionizing radiation experience an early transient incapacitation (ETI) which is characterized by akinesia and decrements in some motivated behaviors (Mickley & Teitelbaum, in press). Typically, performance on a behavioral task is compromised for a period of up to 1/2 hour after irradiation. This ETI is followed by temporary recovery, progressively poorer performance and, in a few days, death. While the ETI phenomenon has been observed often and in a variety of other species (Chaput & Wise, 1970; Young, Chapman, Barnes, Brown, & Hurst, 1968), the etiology of this disorder still remains unclear. One physiological factor contributing to the ETI may be blood pressure, which often drops dramatically during periods of postirradiation performance decrement (Turbyfill, Rouden, & Kieffer, 1972; Brunner, Bogo, & Henderson, 1975).

The search continues for reliable prophylactic drugs which are capable of altering the effects of ionizing radiation. Some success in this regard has been achieved in prolonging the life of x-irradiated mice with injections of estrogens (Patt, Straube, Tyree, Swift, & Smith, 1949; Thompson, Simmons, Crawford, & Severson, 1969; Graham & Graham, 1949; Rooks & Dorfman, 1961; Mirand, Hoffman, Reinhard, & Goltz, 1954) although these animals were exposed to doses much lower than that required to produce an ETI. Indeed, low-dose postirradiation phenomena and ETI phenomena may have little in common.

Since estrogens are critical in development of feminine behaviors (Whalen & Edwards, 1966), we were interested to determine if female rats might differ from males in behavioral responses to radiation levels known to produce an ETI. We also wished to know if the radiation-produced acute hypotension would be present in females to the extent observed in males.

Experiment 1

Method

Subjects. Eight male and eight female Sprague-Dawley rats (200-300 g) were used in this experiment. Rats were housed three to a cage in a room that was illuminated from 7:30 a.m. to 4:00 p.m. daily. Ambient temperature was set at 20°C. Purina Rat Chow and water were continuously available.

Apparatus. All rats were trained to avoid foot shock by leaping a distance of 10 cm from the floor of a cage up onto a retractable leucite ledge. The avoidance apparatus consisted of a Plexiglas chamber 37 x 24 x 37 cm in height with a grid floor that could be electrified with scrambled shock. A signal click preceded the foot shock by 5 sec. The leucite ledge was moved into the cage at the onset of the click and retracted after the foot shock (.5 ma for 10 sec). A microswitch on the ledge indicated if the subject successfully moved up to the shelf before the shock. The intertrial interval was held constant at 30 sec.

Procedure. All subjects were trained to meet an avoidance-response criterion of 100% over 50 consecutive trials. Rats received at least 50 trials once a day until this criterion was met. This usually took about 5 days. On a following day subjects were irradiated, immediately after which 50 trials were again administered in the avoidance chamber.

The Armed Forces Radiobiology Research Institute's Linear Accelerator was used to expose rats to high-energy electron radiation. Each animal was irradiated individually with its right side to the source while in a constraining polyethylene tube. Electrons were accelerated to an energy of 18.1 MeV at a peak beam current of .44 amps. Pulse duration was 4 μsec, and pulses were delivered at a rate of 15/sec. In order to make the field more uniform, a water-scatter device was employed which lowered the energy of the beam by 3.5 MeV. Rats were positioned 3.5 m from the source. Each subject received a midline tissue dose of 10,000 rads (13 rads/pulse) which has been demonstrated to consistently produce an ETI (Mickley & Teitelbaum, in press). Total exposure time was between .8 and 1.2 min. Dosimetry was accomplished by using .05 cc tissue-equivalent ion chambers whose calibration is traceable to the National Bureau of Standards. Animals were retrieved within 1 min after exposure.

Results

The ability of both males and females to avoid shock was drastically attenuated by exposure to 10K rads of high-energy electrons (see Figure 1). However, a Wilcoxon match-pairs signed-ranks test (Siegel, 1956) comparing the two sexes', median percent avoidance at like times (2.5, 5.0, 7.5, 10.0, 12.5, 15 0, 17.5, 20.0 min after exposure) reveals that females avoided significantly more than did the males ($p<.02$).

In addition it appears that the females recovered from the ETI more quickly than did the males. Females returned to their normal performance level of 100% after only 15 min. Males required 5 min longer to reach their preirradiation baseline level.

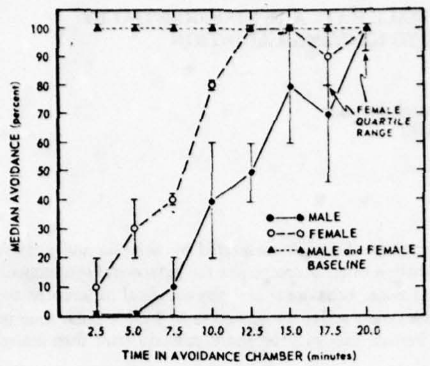


FIGURE 1
Ten thousand rads of high-energy electrons produce an early transient incapacitation in avoidance response. In this period females outperformed and recovered more quickly than males. (Variance measures are the quartile range.)

Experiment 2

Method

Subjects. Eight male and eight female Sprague-Dawley rats (200-300 g) were used in this experiment. The animals were cared for as noted in experiment 1.

Apparatus. Systolic blood pressure was measured indirectly by using Narco Bio-System's occluding tail cuff (32 mm long, 13 mm inside diameter), pneumatic pulse transducer, and programmed electro-sphygmomanometer (ES). Blood pressure readings were obtained while rats were confined in a ventilated transparent plastic animal housing which had a built-in warming element used to raise the environmental temperature of the subject.

Procedure. Each rat was constrained in the animal housing with its tail protruding outside. The metal tubular cuff was slipped onto the tail as far as it would go and then connected to the pump in the ES. Tubing from the pneumatic pulse transducer was taped lengthwise over the animal's tail 3 cm distal to the cuff and then also connected to the ES. The output of the ES was reproduced on a physiograph recorder. Each rat was warmed until a pulse could be detected in the tail. Care was taken to not supply too much heat which would cause behavioral signs of apparent discomfort. When detectable tail circulation had been established, the cuff was inflated to 200 mm Hg at a rate of 20 mm Hg/sec. As the cuff slowly deflated, the systolic pressure was noted as the point at which pulsations first appeared on the descending pressure curve. A baseline systolic pressure reading was taken immediately before irradiation. Subsequent measures were made 5, 10, 15, 20, 25, and 30 min post-irradiation. Irradiation was accomplished as in the first experiment.

Results

Blood pressures of male and female were differentially altered by exposure to 10K rads of high-energy electrons. The median of baseline male blood pressure readings taken immediately before irradiation was 125 mm Hg. As expected, blood pressures of these animals were significantly lowered after exposure to ionizing radiation ($p = .016$, sign test) (Siegel, 1956). (See Figure 2.) Baseline female blood pressure (median =90 mm Hg) was normally lower than for the males. Surprisingly, however, radiation exposure caused a significant increase in the blood pressure of these rats ($p = .016$, *sign test). A comparison between difference scores (the median of individual post-irradiation* pressure minus baseline pressure) computed at different time periods (5, 10, 15, 20, 25, and 30 min) immediately following exposure reveals a significant difference between the blood pressure changes of male and female rats during the ETI period ($p = .001$, Mann, Whitney U Test) (Siegel, 1956).

Discussion

Rats of different sexes exhibited measurable different acute behavioral and blood pressure responses after exposure to 10K rads of high-energy electrons. Female rats outperformed males on an active avoidance task after irradiation. Although females did exhibit a decrement in motivated behavior, it was less pronounced and of shorter duration than that observed in the males. Postirradiation physiological differences between the sexes were also observed. Male blood pressure was consistently lower than baseline for at least 30 min after irradiation. On the other hand, females exhibited a small but significant hypertension in this same period.

In the field of radiobiology there is a sufficiently large number of investigations which indicate a correlation between the radiosensitivity of organisms and their physiological state (Dugger, 1936). These studies, along with the present investigation, are of importance primarily as substantiation of the possibility of active intervention in the organism's reaction to radiation exposure. In other words, data which show a correlation between the radiosensitivity of an organism and its physiological state or sex constitute a certain experimental substantiation of the possibility of finding substances which protect the organism against damaging action of penetrating radiation.

The data presented here demand some underlying physiological mechanism(s) to account for the behavioral differences observed between males and females which offer females a measure of what seems to be natural radioresistance.

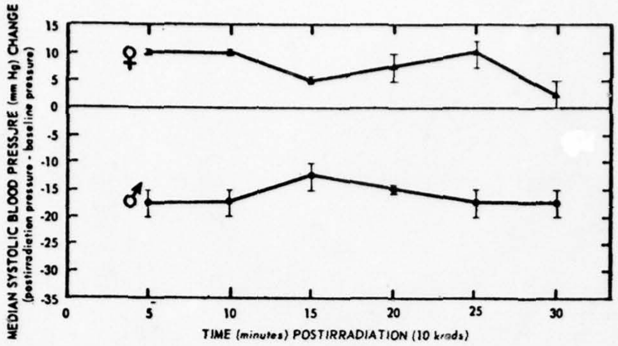


FIGURE 2

Male (♂) systolic blood pressure was reduced (compared to preirradiation baseline) during the 30 min after 10K rads. In the same time period, females (♀) exhibited a relatively mild hypertension. (Variance measures are the quartile range.)

REFERENCES

Brunner, A., Bogo, V., & Henderson, E.A. Dose-rate effects of ^{60}Co irradiation on performance and physiology in monkeys. Topical Report DNA-3660T, Defense Nuclear Agency, Washington, D.C., 1975.

Chaput, R.L., & Wise, D. Miniature pig incapacitation and performance decrement after mixed gamma-neutron irradiation. *Aerospace Medicine*, 1970, *41*, 290-293.

Duggar, B. M. *Biological effects of radiation*. New York: McGraw-Hill, 1936.

Graham, J.B., & Graham, R.M. Pharmacological modification of resistance to radiation - A preliminary report. *Proceedings National Academy of Sciences*, 1949, *35*, 102-106.

Mickley, G.A., & Teitelbaum, H. Persistence of lateral hypothalamic-mediated behaviors after a supralethal dose of ionizing radiation. *Aviation, Space, and Environmental Medicine*, in press.

Mirand, E.A., Hoffman, J.G., Reinhard, M.C., & Goltz, H.L. Sex hormones as protective agents against radiation mortality in mice. *Proceedings of the Society of Experimental Biology and Medicine*, 1954, 86, 24-27.

Patt, H.M., Straube, R.L., Tyree, E.B., Swift, M.N., & Smith, D.E. Influence of estrogens on the acute x-irradiation syndrome. *American Journal of Physiology*, 1949, *159*, 269-280.

Rooks, W.H., & Dorfman, R.I. Estrogen radioprotection in mice. *Endocrinology*, 1961, *68*, 838-843.

Siegel, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

Thompson, J.S., Simmons, E.L., Crawford, M.K., & Severson, C.D. Studies on the mechanisms of estradiol-induced radioprotection. *Radiation Research*, 1969, *40*, 70-84.

Turbyfill, C.L., Roudon, R.M., & Kieffer, V.A. Behavior and physiology of the monkey (*Macaca Mulatta*) following 2500 rads of pulsed mixed gamma-neutron radiation. *Aerospace Medicine*, 1972, *43*, 41-45.

Whalen, R.E., & Edwards, D.A. Sexual reversibility in neonatally castrated male rats. *Jounral of Comparative and Physiological Psychology*, 1966, *62*, 307-310.

Young, R.J., Chapman, P.H., Barnes, D. J., Brown, G.C., & Hurst, C.M. Behavioral and physiological responses of *Macaca Mulatta* monkeys to supra-lethal doses of radiation. SAM-TR-68-73, USAF School of Aerospace Medicine, Brooks AFB, TX, 1968.

INVITED PAPERS[1]

[1]Several papers were submitted by authors although they were not presented at the Symposium.

VALIDITY AND REALIABILITY IN THE EVALUATION OF INSTRUCTIONAL SYSTEMS

John F. Brock
Navy Personnel Research and Development Center

Traditional measures of instructional systems have followed one of several conventional experimental paradigms. In all, test scores within the context of the system are the measures of interest. This paper suggest that these techniques are measuring reliability of instruction, but not validity. A system by which instructional validity can be measured is modeled.

Typically, instructional systems are measured by testing students as they exit the system. Reliability and validity in this context most often refer to the *test* characteristics, not the *system* characteristics. Although books (e.g., Gronlund, 1968) and complete issues of journals (e.g., Tiemann, 1976) have been dedicated to instructional evaluation, confusion between test reliability and validity and instructional reliability and validity still exists.

A test used to measure course effectiveness may be reliable and valid in measuring student performance in a course that is neither. In this paper, reliability of an instructional system will refer to the effectiveness of the system in terms of percentage of students reaching course criteria. Validity will refer to the degree to which the instructional system is meeting the needs of the larger system into which the course graduates feed.

It appears generally agreed upon that the way to evaluate persons within an instructional system is to test their performance on some set of relevant tasks. This same testing procedure typically provides the data for a course effectiveness index. In fact, the Navy's instructional design process has performance test construction as the first major effort, once the tasks for training have been selected (Chief of Naval Education and Training, 1975). Performance testing is the beginning (i.e., the base of all instructional design) and the end of the instructional process (Brock, 1977).

A model for the evaluative process is lacking; particularly, there is a lack of prescription for *where* evaluation of the system fits, *when* evaluation should be done, and *what* should be done with the results of any evaluation.

Brethower and Rummler (1977) propose three general system models which will be discussed in the context of reliability and validity measures of instructional systems. Figure 1 presents these three systems.

The ballistic system is of little concern to the topic of this paper. Since there is no measurement, validity and reliability are either presumed or of no concern. Briefings are an example of a ballistic system in, at least, a quasi-instructional setting.

The guided system fits the more traditional military training system. Assume that the circle with the X in it is the measurement point for system evaluation. At this point, the student is tested and inferences are drawn about the individual student, the aggregate students, and the instructional system. This is the point of the reliability measure. Number of students meeting some predetermined number of objectives has been the traditional quantitative evaluation of this system. The danger of this kind of evaluation will be discussed below.

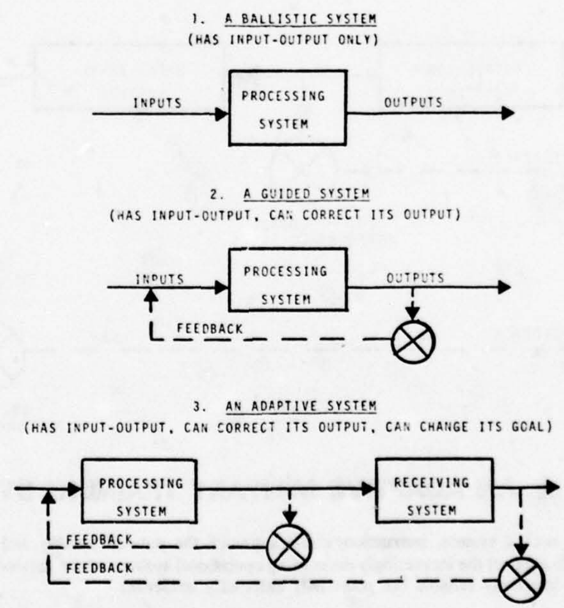


FIGURE 1. GENERAL SYSTEMS MODEL (Brethower and Rummler, 1977)

The adaptive system provides for two points of evaluation; one immediately after training, a second on the job. The first evaluation point measures how well the system objectives are being met; the second point measures how well the instructional system is meeting the needs of the operational system. It is only in an adaptive system, with its appropriate feedback loops, that the instructional system has the necessary information to change its objectives. The second point is, therefore, the point of validity measurement.

The danger in both the guided and adaptive systems is the seductiveness of quantitative data. Let us assume that in a pilot training program there are one hundred objectives. Let us further assume that 95 percent of the pilot students are meeting in excess of ninety of the course objectives. With such a gross measure, no trouble in the system is indicated. However, a careful look at what objectives are not being met could identify symptoms of an ailing instructional program, e.g., 90 percent of the students not being able to eject.

Testing of students at the completion of training only makes sense if a qualitative, as well as a quantitative, analysis of the tests are made. Semple (1974) refers to four levels of instructional system evaluation, based upon work done by Jeantheau (1973). The first three of these levels are reliability measures; the guided and adaptive systems are equally able to be evaluated at these levels.

The first level of evaluation is qualitative. Content, methods, media, and procedures are examined in terms of particular objectives being met or not met. Sources other than performance measures are used; however, the present writer views them as secondary.

The second level evaluation is non-comparative performance measurement. Essentially, this means testing the student's performance before, during, and after training. The degree of improved student performance is, presumably, highly correlated with the quality of the instruction.

The third level involves comparative measurement. Two instructional systems with the same objectives are compared on how well the objectives are met. Typically, this kind of comparison would only be possible for small units of instruction or alternative training devices.

The fourth level is only available with the adaptive model: transfer of training—the comparative measurement of task performance in an operational system. This is *the* test of the validity of an instructional system. It is also the most difficult measure to obtain.

How to measure performance was the recent topic of a symposium sponsored by the Navy Personnel Research and Development Center (Pope and Meister, 1977). Although all three services were represented and most of the leading laborers in the field of performance measurement presented papers, no magic solutions were forthcoming. However, it is clear that all three services are engaged in improving performance measurement techniques (e.g., Anderson and Pickering, 1977; Katz, 1977; Waag and Knoop, 1977).

What is herein proposed is the adaptive model of instructional system evaluation based upon at least two measures of performance: (1) at the completion of the instruction and (2) in the job environment. The first measure is an indicator of the internal state of the system—is the system meeting its objectives? The second measure is, first of all, an indicator of how well the system is meeting the needs of the consumer and, secondly, how much of the instruction is staying with the graduate of the system. Without both measures, information about instructional systems will continue to lack explicitness and, therefore, to provide the feedback necessary for appropriate corrective actions to be taken. Figure 2 is a schematic of this system.

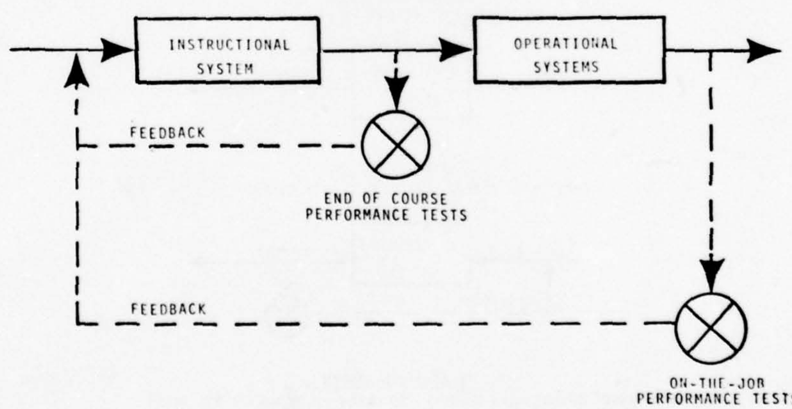


FIGURE 2. AN ADAPTIVE MILITARY TRAINING SYSTEM

If the services could implement such a system, instruction would appraoch the state of validity and reliability necessary to maintain cost-effective instructional systems to support the increasingly demanding operational systems of the services. Without such a system, instructional systems will continue to be internally reliable but potentially externally irrelevant.

REFERENCES

Anderson, A. and Pickering, E. Performance measurement of shipboard operational skills. In L.T. Pope and D. Meister (Eds.) *Symposium Proceedings: Productivity Enhancement: Personnel Performance Assessment in Navy Systems.* San Diego: Navy Personnel Research and Development Center, October 12-14, 1977, 33-54.

Brethower, W. S. and Rummler, G.A. Evaluating training. *Improving Human Performance Quarterly,* 1977, *5,* 3-4; 103-120.

Brock, J.F. Performance testing in instructional systems. In L.T. Pope and D. Meister (Eds.) *Symposium Proceedings: Productivity Enhancement: Personnel Performance Assessment in Navy Systems.* San Diego: Navy Personnel Research and Development Center, October 12-14, 1977, 303-321.

Chief of Naval Education and Training. *A plan for centralized management of instructional systems development within the Naval Education and Training Command.* Pensacola, FL: Chief of Naval Education and Training, 15 August 1975.

Gronlund, N.E. (Ed.) *Readings in measurement and education.* New York: The McMillan Company, 1968.

Jeantheau, C.G. *Handbook for training systems evaluation.* (NAVTRADEVCEN 66-C-0113-2). Orlando, FL: Naval Training Device Center, January 1971.

Katz, M. Planning for performance measurement R&D: U.S. Army. In L.T. Pope and D. Meister (Eds.) *Symposium Preceedings: Productivity Enhancement: Personnel Performance Assessment in Navy Systems.* San Diego: Navy Personnel Research and Development Center, October 12-14, 1977, 361-380.

Pope, L.T. and Meister, D. (Eds.) *Symposium Proceedings: Productivity Enhancement: Personnel Performance Assessment in Navy Systems.* San Diego: Navy Personnel Research and Development Center, October 12-14, 1977.

Semple, C.A. *Guidelines for implementing training effectiveness evaluations* (NAVTRAEQUIPCEN 72-C-02-9-3). Orlando: Naval Training Equipment Center, April 1974.

Tiemann, P.W. An evaluation issue. *Improving Human Performance Quarterly,* Fall-Winter 1976, *5,* 3-5, introduction to a special issue.

Waag, W. and Knoop, P. Planning for aircrew performance measurement R&D: U.S. Air Force. In L.T. Pope and D. Meister (Eds.) *Symposium Proceedings: Productivity Enhancement: Personnel Performance Assessment in Navy Systems.* San Diego: Navy Personnel Research and Development Center, October 12-14, 1977, 381-392.

COMMUNITY PSYCHOLOGY AS A VIABLE DOD ACTIVITY

by

V.P. Luchsinger, Texas Tech University
R.P. Anderson, Texas Tech University
R.F. Hartzell, DOA, Office of Surgeon General
R.H. Rath, Jr., DOA, William Beaumont Army Medical Center, El Paso, Texas
E.G. Bleker, Trinity Valley Mental Health Center, Fort Worth, Texas

Community Psychology is emerging as a recognized field of practice in the behavioral sciences. True inter-disciplinary efforts in the behavioral sciences have been sporadic. Universities have not generally been able to combine the behavioral fields into coherent inter-disciplinary combinations.

Luckily, the world is not organized like universities. Various communities in the western world have captialized upon the potential of combining human resource approaches to the mitigation of human problems, individual and collective. Target populations have included pre-school children, teen-agers, troubled marriage partners, victims of alcohol and drug abuse, retirees, abused wives and children, mental health counselees, and troubled senior citizens.

The beauty of the community psychology approach has been the integrated "systems" approach to the behavioral sciences meeting the needs of human population. In addition to developing the practice of community psychology as an integrated discipline, institutional frameworks are being developed for the practice of community psychology as a part of the total health care delivery system framework. This paper will explore the Department of Defense implication as an example of such an institutional framework.

The Department of Defense Approach

No uniform DOD plan has been formulated to implement a community psychology program. The 1st Annual Training Conference in Community Psychology, sponsored by the Division of Community Psychology, American Psychological Association, has provided program guidelines which have relevance for DOD community programs. Initial efforts to train community psychologists and to gradually implement community psychology practices in the DOD installations that possess comunity characteristics must address the following problems:

1. A lack of qualified personnel to fill positions that require community psychology skills.
2. The lack of experience in identifiable military communities being able to maximize and utilize internal resources and external resources blended together as coordinated primary health care providers in the community psychology sense.
3. The clear cut charter for leadership in implementation of cost-effective innovations in mental health care delivery systems.

Several overseas and continental installation complexes provide a logical test bed for the implementation of community psychology programs from the standpoint of:

1. Location of major medical facilities.
2. Location of major training or troop facilities.
3. Collections of large populations of dependent and or retired personnel.

The goal of community psychology programs on a test basis would provide for the blending of competency-based training and practice directed toward the facilitation of broadly based community development. Identity and consciousness of community concepts are critical to the practice of community psychology. This development would be achieved through the systematic application of empirically derived behavioral and social science principles within naturally occurring socio-historical trends. DOD behavioral scientists and professionals (social workers, psychiatrists, and psychologists) would need to be recruited on the basis of demonstrated competence in basic empirical and clinical skills at the post-graduate level. Professional training objectives would then seek to develop further assessment, intervention, and evaluation skills at all levels of organizational and community hierarchies from individuals to large systems.

Other productive features of installation practice of military community psychology would be the provision of behavioral science consultation, health education for the consumer, preventive mental health education for the community members, as well as consultation for community and military unit leaders.

Implications of the DOD Behavioral Science Professional

This paper will close with a review of the chief implications for the behavioral scientist practicing in the community psychology mode:

1. Professional and personal development would progress to a point where one reaches a clear "scene of community" as a frame of reference.
2. An appreciation for the interdisciplinary nature of community development. The process orientation toward conflict development and resolution would preclude rigid identification with special interest groups.
3. Practice more from preventive-educative mode than from a treatment mode.
4. A firm grasp of empirical methodologies and clinical sensitivities as an essential prerequisite for the broad range of possible community mental health activities.
5. Affirmation of the right of individuals to collectively decide for themselves the future directions of their community. Behaviorally, the community would be assisted to facilitate the development of community internal resources rather than fostering interminable dependency on the "expert".

SITUATIONAL LEADERSHIP FOR WOME[illegible] THE MILITARY: A PROPOSAL

Loren I. K[illegible]
U.S. Na[illegible]

What kind of leadership will be required of and for the increaing numbers of women in the military? Leadership models and research to date have not provided concepts upon which situationally effective leader behavior can be postulated. The design and execution of leadership workshops, seminars, research and other educational and training experiences indicate that women: have been denied experience in leadership roles, because of sex role differences, have a concomitant lack of female leaders as models, and have lower leader position aspirations and realizations than males. The use of a Situational Leadership Influence Model (SLIM) in which the behavior of the leader (LB) is a function of the elements of leader, followers and the situation appear to be particularly applicable to women in the military because the stereotypes of past leader behavior or past leader experience is not a vital factor.

The Defense Officer Personnel Management ACT (DOPMA) provides substantially similar personnel statues and career progression opportunity for all active duty officers below general and flag rank of each service within the Department of Defense. With respect to women, significant changes have been incorporated the effect of which is that basic provisions relating to appointment, promotion accountability and retirement will apply equally to both male and female officers. A notable change in ROTC in recent years is the increase in the number of women enrolled. This statement is typical of those being released by the service academies.

> "Of the 327 persons entering the Coast Guard Academy class of 1980, 38 were young women. They have participated in every phase of cadet training with their male counterparts. They have accepted every academic challenge, participated in every review, sailed on the Eagle, and 'squared corners' in Chase Hall. Their performance has been on par with those of the all-male classes during the Academy's 100-year history. Of the 38 who entered, 28 remain today. That attrition rate compares favorable with the male cadet rate. The Coast Guard expects to see women commissioned in June of 1980 as fully qualified and proud Coast Guard Academy graduates."[1]

The increase of women in the military is well reported, as indeed is the recognition of females in leadership and management positions through the United States. However, there appears to be a strong paradox, between the emphasis and amount of publicity that selected women in both civilian and military life receive and the actual impact they may have in the leadership and management arena.

Number and Percentage of Females in the Military[2]

	Officer	Percent	Enlisted	Percent
U. S. Army	5,420	5.5	45,229	6.7
U. S. Marine Corps	407	2.1	3,423	2.0
U. S. Navy	3,715	0.7	19,202	3.7
U. S. Air Force	5,151	5.2	33,362	7.0
U. S. Coast Guard	64	1.4	605	1.9

Whether the number of women is representative of a start in a comprehensive equal opportunity plan, or whether the number of women is an appropriate recognition of the necessarily important contribution that over 52% of our population can and should make to our national security, or whether the number of women in the military is sheer tokenism is not the point. The important feature is that females have been and are to be increasingly integrated into the military (the degree of which will be the subject of litigation, and Department of Defense and individual service administrative policies)[3]. Are there concepts, theories, plans and programs that can best assist in the accomplishment of immediate goals while providing a firm foundation for true change? What will be the leadership requirements of and for the present and increasing number of women in the military?

The preponderance of leadership research upon which present leadership theory and practice is based, has as its foundation a white, middle-class population. Further, a basis of current theoretical studies is from a male dominate American culture of the 1940-1970. There is a simple question that has not been adequately addressed: Is there a difference in the current and predicted leader follower behavior of males and females?

The traditional sex-role difference claims differences between male and female leader follower behavior. This sex role difference is also paralleled in the early leadership research. Starting with the Great Man approach, i.e., leaders are born, not made, not too suprisingly, these leaders were male offsprings of nobility or other elites. Next in research was the Trait approach—in which leaders can be selected through the identification of Traits. Traits suggested as being important have been those associated with an individual's behavioral characteristics (sociability, aggressiveness, self-confidence, etc.) aptitude (intelligence, originality, judgement, etc.), and biographical profile (employment history, family and educational background, etc.). In recent years, the Trait Approach studies have focused on the use of interest, aptitude and personality tests.[4]. Tests, that for the most part, are also based upon white, middle class populations. Correlation studies between Trait variables and levels of leader effectiveness cannot identify specific leadership traits nor make significant generalizations about leader traits to the general male populations, much less for the actual male female population.

Yet male and female traits (possibly sex role difference?) are frequently cited as the reasons for specific or general male female follower or leader behavior, i.e. "women are too soft", "women are not aggressive." Curiously, this basically false premise of Traits is often used to defend or establish females in leader positions, i.e., "A woman president would have more compassion and improve the quality of life for the nation."

Often, the Trait theory is combined with stereotyped feminine characteristics to show that women are ineffective leaders or followers. In 1971, LCDR Beth F. Coye, (W) USN, in her pioneering study, *The Future of the Restricted, Unrestricted Line Officer*, generally concluded that a posture of women's leaders has been accepted wherein a low profile was maintained. A type of psychological discrimination toward the Wave existed similar to that experienced by professional women. This influenced leader behavior. Waves were continuously conditioned to believe that they were never meant to aspire very high and that inequality of opportunity was part of the natural state of affairs. There was apprenhension about feminine leadership. Leadership by its nature was held to be "aggressive" within societal norms. Female leaders must be both "aggressive" and "feminine", yet these appeared to be an inherent contradiction in terms.[5].

The importance of role models in general personality development is accepted. In current research conducted, both male and female personalities that one might consider recognized leaders are rarely listed. However, even that rarity surpasses the number of times that "mother" has been listed as an admired leader.[6] The difficulties caused by sex role difference to female leadership must be compounded by the lack of female leader role models.

The literature of leadership research continues with the Functional Role Approach—leadership exists in a group when the various tasks and maintenance functions are performed appropriately; the Continuum or Styles of Leadership Approach—leadership can be differentiated into authoritarian or democratic styles; and, the emerging Situational Approach—the leader exercises leader behavior appropriate to the situation. This latter may be a most fruitful approach for women in the military.

It is obvious that there is a basic lack of leadership information regarding female leaders and followers. The issue is important, perhaps vital, to the military. Research to determine if there are female leader follower leadership behavior differences, their extent, possible causes, and rectification—if desired—should continue. In the meantime, it is proposed that situationally appropriate models of leadership be used as the basis for leadership and leadership training for women in the military.

Most empirical and theoretical analysis of leadership attempts to relate leader behavior to various measures of organizational effectiveness. However, leader behavior should be viewed as only one element of leadership. The followers and the situation must also be included. Leaders cannot lead without followers. In any group task situation, followers are vital. The Situational Leadership Influence Model (SLIM) posits a direct correlation between follower influence development (ID) over a period of time, and leader behavior (LB).

Beginning with structuring leader behavior, the appropriate behavior for working with undeveloped followers or groups, the SLIM suggest that the leader behavior should move through (LB-1) predominate structuring leader behavior to (LB-2) high structuring and low consideration leader behavior to (LB-3) equally high levels of structuring and consideration behavior to (LB-4) high consideration, low structuring behavior to (LB-5) equally low consideration and structuring behavior as (and if) the followers influences develops (Figure 1).

Leader Behavior (LB)

Leader —— LB 1 —— LB 2 —— LB 3 —— LB 4 —— LB 5 ——>

Followers —— ID 1 —— ID 2 —— ID 3 —— ID 4 —— ID 5 ——>

Follower Influence Development (ID) Level

FIGURE 1. The Situational Leadership Influence Model (SLIM). Structure and Consideration elements of Leader Behavior (LB). Follower Influence Development (ID) Level (Moore, 1977).

Levels of follower influence development (ID) are provided by dividing a follower influence continuum into five areas: *Initial, low, average, high-average, and* high (simply marked ID-1 through ID-5) for a specific task. When leading followers who have just formed as group (hence "*initial*" stage), a highly structuring leader behavior (LB-1) has the highest probability of success, with followers exhibiting *low* levels of influence development (ID-2) high structuring and low consideration behavior (LB-2) is applicable; for *average* and *high-average* levels (ID-3 and ID-4), equally high structuring and consideration behavior (LB-3) to high consideration behaviors and somewhat lower levels of structuring behavior (LB-4) appear to be most appropriate; and low consideration-low structuring leader behavior (LB-5) has the highest probability of success working with followers who have a high level of influence development (ID-5).

Influence development (ID) level is defined as the degree of achievement motivation, willingness and ability to take responsibility, and task-relevant education and experience of an individual or a group. Additionally, influence development (ID) is congruent with changes in behavior from passive to active, from dependent to independent, from the ability to behave in few ways to the ability to behave in many ways, from subordinate to equal or superordinate positions, and from lack of awareness and control to awareness and control over self or the actions of followers.[7]

Each dimension can be placed on a continuum assuming that attributes and aspects of personality are applicable to the followers as a group, in group theory. At any time, any group can have its degree of development influence (ID) level established within these dimensions.

Follower influence development (ID) level, then, is the degree of follower behavior observed in verbal and vonverbal manifestations of these dimensions. Hundreds of persons who have participated in a variety of task groups in leadership activities have determined the (ID) level of their groups. The analysis of data was conducted by discussions following group task and *post hoc* review of video tapes. Male and female followers in a wide variety of settings, i.e.: operational military commands, leadership training seminars, community colleges faculties, Federal Executive seminars, curriculum development sessions, military and civilian administrative personnel meetings,

degree candidates classes, community action programs, humanistic educational experiences, executive development sessions, etc., evaluated their own behavior and established their own groups influence development (ID) level.

While jargon for different research groups varied, there was general agreement as to the basic concepts.[8] Female members of groups were able to point to specific behaviors that had occurred in the experience and to establish through and with the group, what levels of development a particular follower behavior might indicate. Thus, the "never have been a leader", female was able to learn and develop appropriate leader behavior on the same basis as a male with extensive leader experience.

Females report gains in skills, and feelings of strength and confidence as as a leader and as a follower. "This situational leadership training experience has had considerable impact on my life because my responsibility for conducting groups has increased greatly in recent years and I lacked formal preparation for doing this."[8]

Experience with female leaders both in training situations and *in situ* indicates that the fundamentals of Situational Leadership Influence Model (SLIM) can be internalized relatively easily. Additional functional learning can be achieved by designing the leadership training experience around knowledge or skill areas that are appropriate to the group of leaders being trained.[9]

Experience with female leaders also indicates that women generally have adequate basic leadership skills in structuring and consideration behaviors to be effective. Developing skills or abilities in determining (ID) levels appears to be immediately productive. There is some data (not yet conclusive) that indicates that female leaders may actually have an advantage in that they have not acquired non-functional leader habits. Male leaders in the military, or other institutions, who have been successful because of a chance matching of their Leader Behavior (LB) and the (ID) level of their followers have a difficult time in changing their behavior as the (ID) level of followers or the situation changes.

Females, even with their lack of experience and role models, but with basic leadership skill using the (SLIM) appear to accomplish leader and follower behavioral changes that can enhance leadership for women in the military.[10]

NOTES

1. Admiral Owen W. Siler, Commandant, U.S. Coast Guard: "The Coast Guard, A Continued Need to be Ever Vigilant"; *Sea Power*, Vol. 20, (April 1977).

2. Strength figures obtained 2 August 1977.

3. The Coast Guard assigned female enlisted and four officers to cutters in 1977. The Brookings Defense Analysis project claims approximately 600,000 military enlisted jobs could be filled by women.

4. L. K. Michaelson, "*Leader Orientation, Leader Behavior, Group Effectiveness and Situation Favorability*" Institute for Social Research, University of Michigan, 1972.

5. A longitudinal follow-up is definitely needed. Opposing ancedotal evidence claims both *status quo* and change.

6. "The Determination of Self Perceived LeadeStyle of Females." A University Research Council grant 1977. Professor Loren I. Moore, Ed. D.

7. C Argyris, *Personality and Organization*, (New York: Harper and Row, 1957)

8. The situational approach is readily adaptable to specific participant groups. "A Situational Leadership Model for College Administrators" was presented at the American College Personnel Association 1977 National Convention; "A Situational Leadership Model for Public Administrators" was presented in 1978 for the American Society for Public Administrations; numerous under-graduate, graduate, and post-doctoral courses, seminars and workshops have been conducted by the author over the last several years throughout the United States.

9. A post-siminar evaluation statement of a Female Director of Counseling.

10. To date the rigorous control of variables required for statistical significance is lacking. However, one can state that situational leadership theories and practices have had profound social significance.

READING TRAINING EFFECTS ON THE PERFORMANCE OF ENLISTED STUDENTS IN THE ARMY FINANCE SPECIALIST COURSE (MOS 73C20)

By

James L. Sherrill
US Army Institute of Administration
US Army Administration Center
Fort Benjamin Harrison, Indiana

June 1977

PURPOSE

To evaluate the worth of reading training administered to poor readers prior to course enrollment.

PROCEDURE

Poor readers assigned to the Finance Specialist Course were identified by a reading ability test given the first day of training. The High School Reading Test (Form W) of the Stanford Achievement Test, Harcourt, Brace and World, Inc., 1965, was used for testing the first 34 students used in the experiment. The Reading Test of the Stanford Task Test of Academic Skills, High School Level (Level 1, Form A), same publisher, 1973, was used for testing the remaining 14 students used. Ninth grade norms were used in both cases in converting the raw scores to percentile scores.

From two to six poor readers were selected from each of 12 Finance Specialist Course classes beginning with class number 3, 27 October 1976 and ending with class number 16 which began 23 February 1977. The median percntile score of all students used in the experiment was nine. These identified poor readers were formed into pairs (each member of each pair having apporximately the same reading ability as measured by the reading test). By toss of a coin, one member of each pair was assigned reading training and the other pair member continued as a student in the Finance Specialist Class in which enrolled. Those students *not* assigned to reading training were known only to the experimenter.

Reading training was conducted four hours each day in the morning by a professionally trained reading instructor licensed by the State of Indiana to serve as the Director of Reading Training for an entire school system, a Reading Specialist, a Reading Consultant, and a Reading Supervisor. In the afternoon, those students assigned reading training supported light duty requirements of the US Army Administration Center. Duration of reading training was from two to six weeks dependent on progress. Release from reading training was timed to coincide with the start of a new class of the Finance Specialist Course. On release from reading training, and regardless of their progress in reading training, individuals were enrolled in the Finance Specialist Course.

RESULTS

There were no course performance differences among those students administered reading training and those not administered reading training. (See Table 1).

Significantly more training time was invested in those assigned reading training than was invested in those not assigned reading training. (See Table 2).

The course performance of a randomly selected sample of 26 good readers was clearly superior to that of the poor readers. All 26 passed. Only one of these students was recycled.

DISCUSSION

No attempt was made to evaluate the language background of those students used in this study. Inspection of the surnames suggest that non-native language backgrounds may have been a significant factor in course performance. Twenty-one of the 48 students (both experimental and control) experienced problems (failed or recycled). Fifteen of these students had surnames which suggest non-native language backgrounds. But, the course performance of the students with non-native surnames was not appreciably different from those students with native surnames.

The listening and speaking skill of the non-native language background students may have been a significant factor in course performance. However, the test used to assess reading skills did not purport to assess listening and speaking skills.

Other uncontrolled factors were present in this experiment which could have had significant effects on the results: interaction between students and between students and instructors, different instructors in the several classes from which the experimental students were drawn, and others. In spite of these possibly real limitations, the results strongly support the need to formally evaluate any reading remediation program before it is adopted as a routine practice.

The relationship between measured reading ability at course entry and the criterion of pass and fail has ranged from .33 to .57 within the Finance Specialist Course. These were point-biserial correlation coefficients.

Of interest, six of the students assigned reading training scored at, or below the first percentile reading score. Of these, three passed outright, one failed after being recycled, and one failed outright. Even the most deficient readers, as measured by the test will frequently do very well in the Finance Specialist Course.

CONCLUSIONS

Reading ability is clearly a significant, but low, factor in poor student performance with the Finance Specialist Course.
Reading training, as administered, had no effect on course performance.
Recycling failing students was less costly than reading training terms of total days invested.

TABLE 1
DISPOSITION OF STUDENTS

	BY FOUR CATEGORIES		RECYCLE AND		BY TWO CATEGORIES*	
	PASS	FAIL	PASS	FAIL	PASS	FAIL
Reading Tng	17	4	1	2	18	6
No Reading	10	4	7	3	17	7

$* X^2 = .11$

TABLE 2
TRAINING TIME INVESTED

	AVERAGE TRAINING DAYS*		
	READING	COURSE	TOTAL
Those Assigned Reading Training	41.4	44.0	78.9
Those Not Assigned Reading Training		47.2	47.2

$* t_{dep} = 8.53 \ (P < .01)$

TACTICAL DECISION MAKING: A LEADERSHIP TRAINING SUCCESS STORY

B. Don Sullivan and Arthur B. Sulkin

BACKGROUND

"The Purpose of the USAF Academy is to Prepare Cadets to be Professional Officers in the United States Air Force"

To this end, an integral phase of the curriculum of the US Air Force Academy is leadership—developing leadership skills in a role as a follower the first two years and as a leader for the last two years. As a part of this training, all basic cadets entering the US Air Force Academy during Summer, 1976, were provided leadership theory and laboratory experiences. This training was the initial step toward preparing these men and, for the first time this year, women for service as Air Force Officers a process which will continue over the four years of their enrollment in the Air Force Academy. The purpose of this paper is to describe the procedures followed in providing a portion of this training and to illustrate its applicability as a learning vehicle in other precommissioning programs and in active duty or reserve Air Force units.

The leadership training during the initial summer Basic Cadet Training period (BCT I) consists of an orientation to Functional Leadership (discussed later in this article) several structured experiences related to the achievement of specified objectives on the Leadership Reaction Course; and concentration on building a "team spirit" within the cadet element. These structured experiences have easily definable goals and a clear measure of success and thus build confidence in the cadets' leadership ability. This training is, of course, but a fraction of rigorous physical and military program experienced by the basic cadet during BCT.

Following BCT I, the basic cadets march to the Jack's Valley Training Area and begin the second phase of the basic training (BCT II). This Phase encompasses combatives, small element leadership, rifle ranges and, among others, the Tactical Decisions Course (TDC) upon which this article is based.

During previous summers the TDC had been conducted in a simulated combat environment with attendent explosive simulators, blank rounds and emphasis on small unit tactics. As a result of this emphasis, the basic cadets tended to concentrate upon combat movement of their units rather than analysis of the situation, plans of action and other essential leedership activities. Even though the excitement of the envrironment (explosions, machine gun fire, snipers, etc.) was considered to be motivational, it detracted from the learning objectives of the course.

Another limitation of the old concept of leadership instruction was the emphasis on solving a puzzle (tasks) with a precise set of rules and resources. This approach was appropriate for the cadet's initial exposure during BCT I; however, since learning was to be a building process, the format was considered too simple. Since a leader seldom faces a situation where the task, resources and restriction are precisely mapped out for him, the TDC was modified to reproduce some common leadership problems, e.g., incomplete inaccurate information; decisions.

DISCUSSION

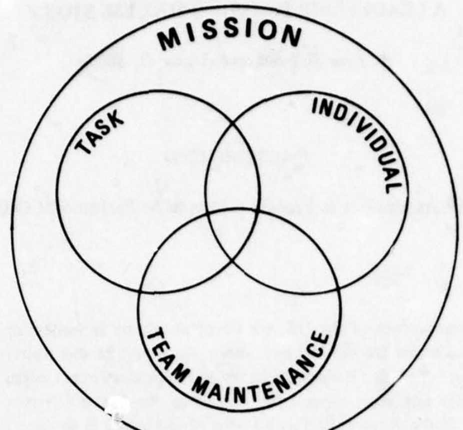
The TDC was modified to accomplish the following primary objective:

"Provide basic cadets with the opportunity to develop and practice decision making skills in a leadership environment. They should learn to cope with the pressures and problems of leadership. As the cadets attempt to accomplish a series of tasks as the leader or as a member of a team, he or she should build the foundation for leadership traits and skills which will be used throughout their Air Force careers."

To accomplish this objective, several new methods were implemented. First, the basic cadets, were given a complete reorientation to the theory of Functional Leadership upon which all scenarios employed on the course were based. This approach to leadership was essentially drawn from the Royal Air Force but is equally applicable to active and reserve Air Force components.

Functional Leadership was considered the best approach to teach cadets because most Air Force leadership takes place within organized groups whose leader and members are clearly identified and differentiated as to their responsibilities for the task of achieving a common goal. Additionally, the Air Force leader usually has legal and or ethical responsibility for his organization, therefore other approaches such as situational, interactional and traitist were not considered adequate for unique military structure. However, this formal military organization creates side effects which were highlighted for the cadets. Organizational principles such as task specialization, unity of command, chain of command, and span of control were emphasized as sometimes frustrating to an Air Force cadet or officer. Since this system often creates situations in which members are dependent, subordinate and passive to the leader for their rewards, penalties, directions,and communications the dynamic and aggressive cadet may meet with frustration. This relationship with its potential pitfalls between the leader and the members was emphasized.

Briefly, the Functional Leadership approach recognized three overlapping major components or needs (task, team maintenance, individual) within the overall mission with designated functions under each need. Schematically, this is shown:



Specific elements of this model include:

TASK FUNCTIONS

Appriciating/Reappreciating the Situation

- **Define the task and determine how it relates to overall mission**
- **Reason out the relevant factors**
- **Understand the limitations**
- **Decide on the best course of action**

Make/Adjust the Plan

- **Utilize previous experience and that of specialists in organization to assist in planning**
- **Assess resources required versus resources availble**
- Assess/Reassess the priorities
- **Provide for contingencies**

Allocate Work/Resources

- **Assign tasks and allocate resources**
- **Assign responsibility and authority for task accomplishment**

Control Quality/Tempo of Work and Performance

- **Check standards**
- **Influence tempo**
- **Ensure coordination**
- Measure effort/time/resources used

Ensure Communication

- Give concise, clear information/instructions
- **Check for understanding**

Set the Standards

Set the example

Correct faults

TEAM MAINTENANCE FUNCTIONS

Instill/Maintain Discipline

Insist on obedience

Punish

Build Team Spirit

Include the group in decision making

Enc Encourage suggestions/cooperation

Motivate, Praise, Give a Sense of Purpose

Assess morale

Know names, faces, background

Reinforce success

Accord Status

Develop potential leaders

Appoint/promote subleaders

Reward effort publically

Train the Team

Encourage education and training

Practice for tasks

INDIVIDUAL FUNCTIONS

Recognize and Use Individual Abilities

Assess ability

Delegate responsibility

Attend to Personnel Problems

Check working and living conditions

Deal rapidly with welfare problems

The basic cadets were carefully briefed that the relative importance of the needs may vary from leadership situation to situation, and, in fact almost always overlap. It was also emphasized that the overall mission is paramount to any of the particular needs (task, team maintenance, individual).

After being briefed, the basic cadets were then rotated, in Elements of ten to twelve persons, through eight stations. At each station the Element was presented with a different scenario each requiring action on behalf of the leader which was related to the various aspects of the Functional Leadership paradigm. In some situations, the leader was required to use a pure democratic or autocratic style of leadership. Generally, each scenario consisted of a "real Air Force" situation (downed B52 crew; Special Forces Team, etc.) in which elements of information, both essential and nonessential, special limitations of terrain or enemy activity and role playing group members all interjected factors which had to be processed and acted upon, as appropriate, by the leader of the group. Specifically, individual cards containing unique information were presented to the leader and three or four element members. Sample scenarios can be seen as *Figure 1 & 2.*

FIGURE 1
"HELICOPTER RESCUE"

1. *LEADER'S INFORMATION*

You are the commander of a downed C-141 crew. You have a rendezvous with a SAR helicopter in a clearing ½ mile to the N.E. of your present position. You are to secure this position and wait for the chopper's arrival.

- Wind is 20 kts. from the N.W.
- Temp: 85°F
- Altimeter: 28.92
- Enemy activity reported in the area

2. *ELEMENT MEMBER'S INFORMATION* (This information is fed separately to randomly selected members of the element. Do not let them discuss their assignments.)

a. You have been previously briefed that the only landing zone for helicopters in the general area is located about 2 miles to the S.E. of your present location.

b. You have been provided with a very important portion of information pertaining to your task. . . . There is a barbed wire barricade North of your position running parallel to the best proposed route to the destination.

c. **** REMEMBER THAT YOU ARE ROLE-PLAYING.

You are to play the part of a very argumentative crew-member. Try to find fault with most of the plans that your acting leader proposes.

d. You recall that someone told you that the enemy moved out of this area at least two days ago. In fact, you are certain that this information was provided you during an intelligence briefing.

FIGURE 2
"MISSING MAN"

1. *LEADER'S INFORMATION*

You are the commander of a downed aircraft crew. You have gathered your crew and one of the men is missing (the tail-gunner). It could be days before you are rescued or can walk out. There is a friendly encampment located 3 miles to the N.W. of your position.

- Wind is 0-5 kts. from the N.E.
- Temp: 39°F
- There is heavy anti-aircraft fire in the area
- The 82nd Airborne has just landed 10 miles to your S.E.

NOTE: YOU *MUST* USE A PURE AUTOCRATIC STYLE

2. *ELEMENT MEMBER'S INFORMATION*

a. You remember from your trip down in the parachute, that the tail gunner was going down in an area about a mile to the north northwest

b. From the preflight map you have seen you estimate your position to be deep inside enemy territory. Those were the most up-to-date maps.

c. You bailed out of the plane right behind the tail gunner, and he was shot up pretty bad. He wasn't looking any too good the last time you saw him.

d. You are the aircraft navigator. According to your calculated position before being shot down, you crossed over into firendly territory 20 minutes before you were shot down.

e. **** REMEMBER THAT YOU ARE ROLE-PLAYING!

You are playing a cowardly role. You can't go after the other guy, after all you could all be killed. He was wounded, he's probably dead anyway. There is no sense in all of you getting killed for a corpse.

The cadets were then evaluated and critiqued by the scenario instructor (cadet upperclassmen) who used as his criteria the principles set forth in the Functional Leadership model. The basic cadets were carefully briefed that there was no "school solution", rather, their success would be judged on the process of their decision making and their adherence to the Functional Leadership principles.

CONCLUSIONS

Although the Functional Leadership model and the exercise of military leadership in the field was relatively new to the majority of basic cadets, the maturity and judgement which they demonstrated in making decisions affecting the welfare of their Element was impressive. For the most part, they adapted the model and were able to discern the relative importance of the needs in order to accomplish the mission and to react to changing needs as various factors were introduced in the scenarios.

Certain of the basic cadets expressed dismay that there was no "school solution" provided for the various exercises. This observation was utilized by the instructors to emphasize that since leadership is both personal and situational, a given solution is not possible but that use of the model will provide a basic guideline to systematically evaluate the altervatives which are available to every leader.

As noted earlier, this was the first year for women at the Academy. The female basic cadets were selected for leadership roles on even par with males. They took firm control of the situation and processed cogent data equally well as the male basic cadets. Their instructions and leadership styles were well received by their fellow team members. The only criticim of the female leaders was that in some instances their voices could not be heard over the rattle of underbrush and the crunching of pinecones.

Overall, the morale of the teams completing the course was very high since all decisions made were those of team members and thus responsibility was clearly fixed within each team.

At the conclusion of the course an anonymous course survey (Figure 3) was completed by each basic cadet. Those evaluation categories which related to course content are summarized below:

FIGURE 3
COURSE INSTRUCTOR CRITIQUE

PLEASE ANSWER THESE QUESTIONS OBJECTIVELY AND POINT OUT AREAS IN WHICH THIS COURSE AND YOUR INSTRUCTOR CAN IMPROVE. THIS CRITIQUE WILL NOT AFFECT YOUR RATING IN ANY WAY, NOR WILL IT AFFECT YOUR SQUADRON'S RATING OR YOUR INSTRUCTOR'S RATING.

COURSE

	Excellent	Good	Satisfactory	Weak	Unsatisfactory
1. How would you rate the course as a whole?					
2. How would you rate the eight leadership exercises?					

3. Would like more of this type of training?

____

4. What improvements would you suggest for the course? (Be as specific as possible)

____

INSTRUCTORS

TO BE OF MAXIMUM ASSISTANCE, PLEASE SPECIFY INSTRUCTOR'S NAME OR AREA OF TRAINING ON THE FOLLOWING QUESTIONS:

5. *Did you like the way they presented the material?*

YES NO COMMENTS:

____

6. Did you notice any distracting mannerisms in their teaching?

YES NO COMMENTS:

____

7. Do you think they knew the subject well?

YES NO COMMENTS:

____

8. Do you think they were sincere and enthusiastic? YES NO

9. Do you have any other comments for them?

____

____

a. **Overall course value**

Excellent	45%
Good	55%

b. **Overall leadership scenarios (8) evaluation**

Excellent	55%
Good	35%
Satisfactory	8%
Weak	2%

c. **General comments - 35% of the basic cadets indicated that they wanted more realism in the course.**

The Functional Leadership model and the operational training to reinforce this model, as described above, can well serve as a useful tool in active and reserve Air Force units desiring to improve the leadership training and capabilities of their personnel. The scenarios, of course, can be modified to apply to a variety of situations most suitable to the unit's mission, e.g., administrative, logistics or medical, since the Functional Leadership model is equally applicable to both tactical and managerial missions. While this particular course was oriented toward emerging leaders in the form of basic cadets, the model and the TDC are equally applicable to training experienced leader personnel interested in developing a new approach to leadership. The administrative overhead to implement the course is minimal since each station consists of only a knowledgeable evaluator, the situation cards and a scenario evaluation form. Although an outdoor setting lends realism to tactical decision making, it certainly is not an absolute prerequisite to success since the emphasis is upon process rather than accomplishing a particular task such as taking of a hill, skirting of a machine gun nets, etc.

CCAF: A SPACE AGE APPROACH TO OCCUPATIONAL PROGRAM DEVELOPMENT

William A. Wojciechowski
Community College of the Air Force

OVERVIEW: Education program development within the Community College of the Air Force (CCAF) is consistent with its philosophy and in line with trends in both Air Force and civilian occupational education. CCAF intends to focus the educational effort of airmen toward obtaining a high quality, occupationally oriented, career education which will enable them to serve as a master of and a supervisor in their specialty, either in the Air Force or in civilian life.

The U.S. Office of Education estimates that 80% of the work force will require occupational skills in the 1980's as contrasted to baccalaureate skills. Concurrent with this changing emphasis on needed educational skills, we have seen the Air Force enlisted force decline from more than 800,000 personnel to slightly more than 500,000 in less than eight years. With this reduction has come an expansion of technology requiring our personnel to be familiar with more complex weapon systems. Moreover, we see a growing need for more sophisticated supervisory skills. Past incidents of notable unrest indicate that knowledge of technology alone is inadequate—skill in human relations acquired through an understanding of the social sciences and the humanities is necessary. Therefore, the Community College of the Air Force has systematically plotted the career education of Air Force enlisted personnel to meet the demands of changing technologies and to develop the social awareness necessary for the effective management of people.

OCCUPATIONAL REQUIREMENTS: Requirements for specific jobs in the Air Force are determined from a review of programming documents, such as the Personnel Manpower Change Program series, which indicate numbers of personnel required in specific occupational specialties. The documents also reflect how many airmen, and with what training, leave the Air Force annually. The quality of airmen needed is indicated in Air Force Manual 39-1, *Airman Classification Manual*, and through means such as job inventory programs conducted by the Occupational Measurement Center of the Air Training Command, the Air Force Human Resources Laboratory, and feedback from using agencies.

Since it has been estimated by the Aerospace Education Foundation and the Air Force Human Resources Laboratory that from 60% to 90% of Air Force occupations have civilian counterparts, we believe a review of civilian occupational requirements is necessary as part of our program development process. The Department of Labor's *Occupational Outlook Handbook* and comparable state guides are reviewed to determine projected requirements for specific occupations. Long range projections are provided by the Department of Labor's Industry-Occupational Matrices and the President's manpower report.

An assessment of the quality of education required is then made through a review of guidelines such as those found in the USOE Technical Education Program series and *Occupational Criteria and Preparatory Curriculum Plans in Technical Education Programs*. Curriculum guides such as those for aviation and electro-mechanical careers published by the American Association of Community and Junior Colleges are also beneficial. The ASEE* Engineering Technology Education Study has been useful as a means of defining parameters of education for technicians. Guidelines prepared by other agencies such as the International Association of Firefighters and the Texas Commission on Law Enforcement Education have been studied and incorporated wherever possible to provide the necessary prerequisites for future licensing and certification of CCAF graduates.

RESOURCES: *Prior to the existence of the CCAF's educational programs* for enlisted persons, little emphasis was given to tying together service provided instruction and the related technical and general education available to the airman through a variety of sources. The Air Force provides technical training via some 2,500 to 4,000 technical training courses to approximately 300,000 students per year. Approximately 80,000 students attend resident technical courses upon entering the Air Force. In addition, airmen receive management instruction through a series of NCO Academies and specialized instruction from agencies such as the School of Aerospace Medicine. Also, there is an extensive system for providing work experience combined with on-job instruction through a dual-channel OJT program—work and study. Completion of this period of apprenticeship is documented attesting to an individual's ability level to *perform* skills necessary in a particular specialty. This supervised training is identified as CCAF's Internship program for which credit is awarded.

Most of the aforementioned instruction is specifically designed to prepare airmen either as technicians or as managers leaders. CCAF integrates this instruction with courses in related general education available from almost 400 civilian institutions which are associated with our 172 Base Education Services Centers. Included in CCAF's programs as an option is limited credit by examination offered through DANTES (Formerly the Armed Forces Insitute (USAFI).

CCAF CONTROL: If these myriad forms of instruction are to be focused toward career relevant education for airmen, program control must be exercised by a central agency. CCAF serves this function for the Air Force. The Program Development Division of CCAF is charged with the responsibility for analyzing service instruction to determine which parts of Air Force instruction:

- Are at civilian post-secondary level. (This evaluation is made by program administrators, analysts, and department heads who have an average of 18 years experience in their occupational specialties as well as undergraduate and graduate degrees).

- Have civilian applicability and or are occupationally related including that which is exclusively Air Force oriented. Subject matter which meets these basic criteria are evaluated on the basis of an average of 30 contact hours of instruction being equivalent to one semester hour. CCAF credit is applicable toward its Associate in Applied Science (AAS) Degree.

Using service instruction as a core and guidelines established by USOE and other standard setting agencies for 2-year occupationally oriented associate level programs, a basic career pattern for the AAS was established. Within a total minimum length of 64 semester hours, a sub-minima of 24 semester hours technical education related to an airman's Air Force occupation, 25 semester hours of general education designed for personal enrichment and to enhance supervisory skills, and 6 semester hours of management instruction make up the basic program pattern. Many technical electives and general education courses in the program are obtained from accredited institutions

---

*American Society for Engineering Education

by the airman during his off-duty time, either with the Air Force providing 75 percnt tuition assistance or with VA assistance. Currently, only 24 hours of credit by examination may be used toward this requirement.

As is clear from the foregoing, we are interested in focusing students' efforts on education related to their occupational specialties within the Air Force in order to maximize job performance and to enhance opportunities for post service employment. Therefore, after outlining 64 hour minimum program patterns, the Program Development Division has reviewed all Air Force occupational specialties and clustred them into 85 programs of study in five major career areas, e.g., in Management and Logistics there are programs in general business, business management, and computer science. In the electro-mechanical career areas, some 17 programs were developed. Each of these was constructed using the best available curriculum guidance. For example, U.S. Office of Education guidance for technical education indicates that programs should provide students with a facility in mathematics, physical science principles relative to technical skills, and ability in communications skills along with the knowledge of a particular occupational specialty. To this core we have added a facility in social sciences, humanities, and management to insure that airmen attaining supervisory positions develop a social awareness necessary to coping with the increasingly complex demands on personnel in management positions.

After the 55 programs were initially developed, they were reviewed by the CCAF Policy Council (annual review), chaired by the Dean and consisting of members of the College staff to include experts in the areas being reviewed. The programs were reviewed against criteria such as the following:

- Does the program provide for Air Force occupational needs?
- Does the program have a civilian occupational orientation?
- Does the program meet the established minimum of 64 semester hours?
- Does the program satisfy minimum criteria for accreditation such as those expressed in the Southern Association's Commission on Colleges guidelines?

After modifications suggested by the Policy Council to the programs are made, advisory panels of consultants are brought into action.

ADVISORY PANELS: To provide assurance that programs developed and controlled by CCAF are consistent with Air Force needs and, where possible, civilian requirements, we have identified advisory panels of consultants. They are representatives from the technical schools which offer instruction reflected in the programs; representatives from the Air Force Training Command Technical Training Directorate of Surgeon General; and representatives from business, education or industry. Wherever possible, representatives from appropriate professional organizations are asked to review CCAF programs and comment on them in terms of adequacy in preparing individuals to fulfill professional duties at the technician level. If licensing or certifying agencies exist for an occupational specialty, members of those agencies are also asked to review appropriate programs to determine their adequacy in preparing students for certification and licensing. Finally, in those areas where employment entry is dominated by unions or specific industries, members of the unions or industry may also be asked to comment. Consultants representing the groups noted above are employed and formally review CCAF programs.

Subsequent to this annual review, programs are modified as necessary by the Program Development Division and formally approved by the Policy Council. Thereafter, the Programs are incorporated into the CCAF catalog for distribution to Air Force schools and bases, recruiters, high schools, Education Services Officers, college registrars, and requesting employers.

ASSOCIATED IN APPLIED SCIENCE DEGREE APPROVAL: When the airman's documentation is received at CCAF, it is reviewed by program administrators to insure that courses completed are from accredited institutions, are consistent with program objectives, and that the airman has a comprehensive grasp of his occupational specialty. Airmen who fulfill AAS requirements are recommended to Policy Council for approval and award of the Associate in Applied Science degree.

FOLLOW-UP: To assure that the degree programs continue to meet the needs of the Air Force and, where possible, civilian employers, follow-up studies are conducted by the Institutional Research Branch of the CCAF Registrar. The studies determine how CCAF programs contribute to the improvement of NCO quality in terms of vocational skills and supervisory competence and how useful the programs are to employers and to other agencies such as colleges and universities. Feedback from these studies, as well as informal feedback from registrars, surveys, and other sources, enable us to modify and improve the programs as necessary.

TESTING THE MASLOVIAN HIERARCHICAL NEED THEORY IN AN AIR FORCE ORGANIZATION

G.C. Saul Young
Air Force Institute of Technology

This study reports on the results of an empirical test of Maslow's need theory. The survey instrument employed four different measures of need strength for each of the Maslovian need types. The target population was a random sample of civilian and military personnel at Maxwell Air Force Base, Alabama. The results showed mixed support for the Maslovian theory.

Introduction

This paper reports on the results of a survey-based test of several mathematical versions of the Maslovian hierarchical theory of Human needs. The mathematical formulations of Maslow's theory were developed by Young (1975). The survey used to measure the need strengths necessary to test the theory is based upon a ten-item instrument developed and validated by Mitchell and Moudgill (1976).

The Maslovian Theory

The Maslovian hierarchical need theory hypothesizes that people operate primarily in one of five distinct hierarchically related types of needs; Physiological, Security, Belongingness, Esteem, and Self-Actualization. Furthermore, Maslow hypothesizes that people move from one need to the next higher need only after the next lower need has become largely satisfied. The relationship between the strengths of these different needs over the average person's lifetime is portrayed in Figure 1.

The main hypotheses flowing from this graphical model of Maslow's theory are the following:

1. For individuals who are clearly operating at a given need level (as indicated in Figure 1 by strength in the "Potent" range), all the lower need levels should be largely satisfied, and should thus display positive correlations amongst each other.

2. Similarly, all the need levels higher than the one at which the individual is clearly operating should have prepotent need strengths, thus all being low and hence showing positive (and strong) correlations among themselves.

3. Finally, the low, prepotent needs should begin to display increasing strength as they begin emerging into potency; and the low, largely satisfied needs should show low, even declining strength as they remain largely satisfied and become even moreso. Therefore, we should see negative correlations between prepotent need strengths and largely satisfied need strengths.

There are numerous other hypotheses that can be posited, including those involving *levels* of need strength (as already implied in the earlier hypotheses); and the type of curve (concave) which should fit the strength of adjacent need pairs, given that one of them is at the clearly operating level of strength. However, space limitations do not allow further development of these hypotheses.

Measuring the Maslovian Need Strengths

Various instruments have been used to measure the strengths of the five Maslovian need types. These stretch from the Porter's work in the early 1960s (1961) to the most recent efforts by Mitchess and Moudgill (1976). Some of these instruments measure importance of a need; some measure degree of satisfaction (or dissatisfaction) of a need, others measure need deficiency (amount-one-wants minus amount-there-is); and others measure strength of the need.

The survey used in this work incorporated all four of these measures, as applied to each of the four upper Maslovian need types. In addition, the survey measured a fifth need type, Autonomy, posited by Porter as falling between Esteem and Self-Actualization.

The Target Population

The survey was administered to a cross-section of the military and civilian workforce at Maxwell Air Force Base, Alabama in September 1977. A total of 850 surveys were distributed; 440 usable responses were received, for a response rate of 52 percent. There was good representation across a number of different job types, racial classes, and rank.

Validating the Measurement Instrument

Correlations were calculated among the four different need strength measures for all five need types. High positive correlations were found between the measure of Need Strength and Need Importance, ranging from +.76 to +.81 across the five need types. Reasonably high positive correlations also showed up between Need Deficiency and Degree of Dissatisfaction, at least for the four highest need types. These correlations ranged from +.53 to +.76. Unfortunately, the other need measure correlations did not indicate congruence among all need measure pairs. In view of this, all need measures were used separately to test the hypothesized relationships.

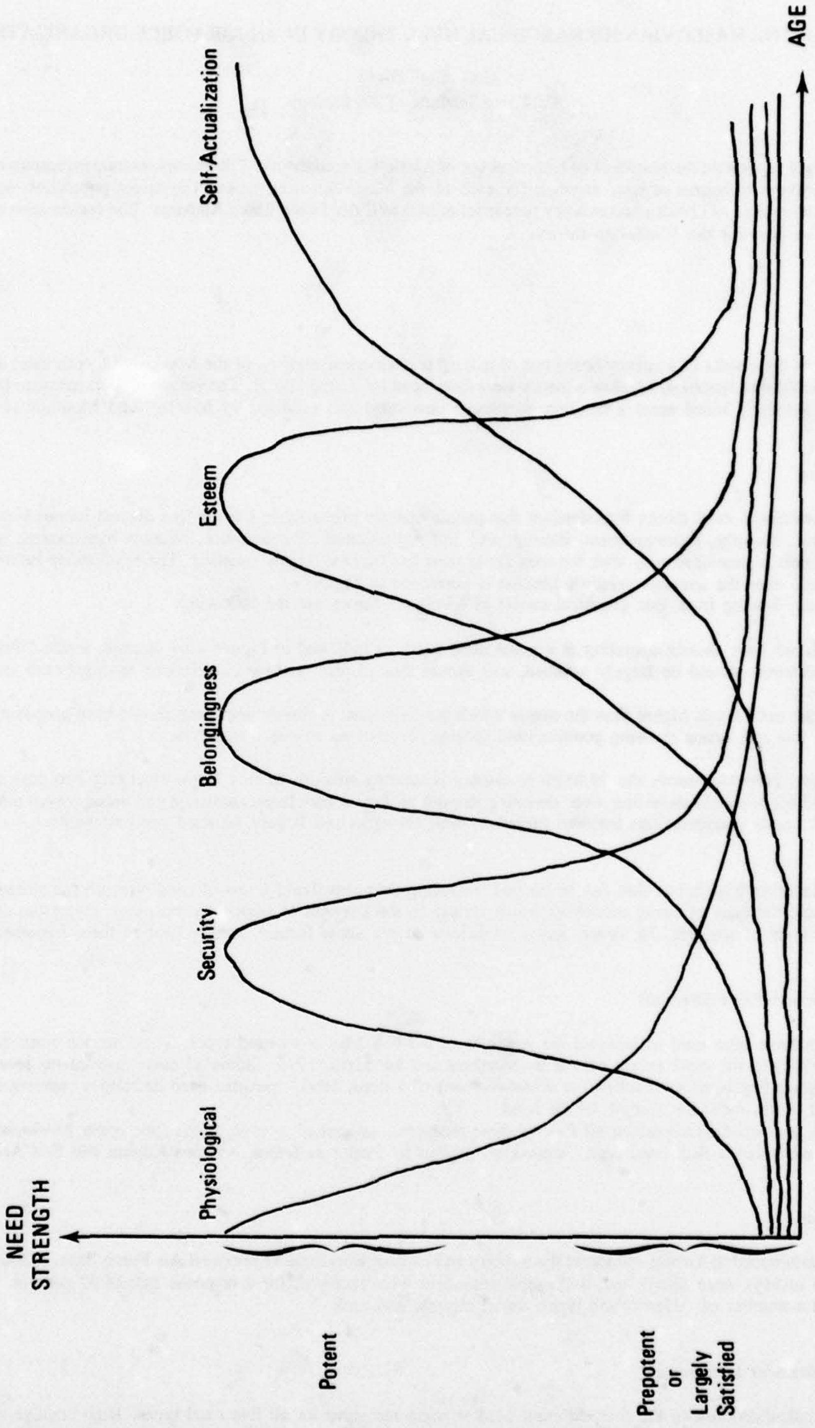


FIGURE 1
Strengths of the Maslovian Need Types as a Function of Age

Results

Table 1 shows results which are representative of all four of the need strength measurements across all five of the need types. This table is based upon the Deficiency measurement of Need Strength (amount-one-wants minus amount-there is). Further, the data used is only for those individuals who showed potent strength (a high Deficiency measure) for their Esteem need. The table displays the need Deficiency cross-correlations among all possible pairs of prepotent (Autonomy and Self-Actualization) and largely satisfied (Security and Belongingness) need types.

Table 1
CROSS-CORRELATION BETWEEN PREPOTENT AND LARGELY SATISFIED NEED STRENGTHS

	Security	Belongingness	Esteem	Automony	Self Actualization
Belongingness	.2601**				
Esteem	.1573*	.3120	1.00**	.4850**	.5324**
Autonomy	.1684*	.4623**			
Self-Actualization	.0968	.3500**		.5519**	

*p < .05
**p < .001
Note: Includes total population - all subgroups and job types.

The boxed correlations in Table 1 should, by Hypothesis 3, be negative and significant; all remaining correlations, by Hypotheses 1 and 2, shoule be positive and significant.

As can be seen, though these correlations are statistically signifiant at better than the .05 level in all but one case, all of them are positive. Thus, Hypotheses 1 and 2 are supported, and Hypothesis 3 is not supported.

This pattern repeated itself for the other three measures of need strength across the other four need types when set at potent strength levels.

Discussion

The survey employed exhausts the possible ways of measuring need strength, but gave consistent results no matter which of the four measures were used. These results were mixed, supporting some portion of the Maslovian theory in two of the hypotheses, but not supporting that theory in the third. What can we make of this?

First, we must remind ourselves that this data was collected within, and refers only to, the working environment. Maslow was not so parochial in defining the arena within which he stated his theory applied—namely, to *all* of the normal life space. Perhaps then, these mixed results are due to the limited job arena of our data collection, or perhaps they suggest, as have other works, that Maslow's hierarchy theory is either not testable empirically, or unlikely to be validated by empirical data. With such mixed results, all we can do at this point is to suggest that additional research be conducted on the issue, especially in more exhaustively developing and testing other hypotheses which seem to be suggested by his theory.

REFERENCES

Mitchell, V. F. and P. Moudgill, "Measurement of Maslow's Need Hierarchy." *Organizational Behavior and Human Performance*, 1976, *16*, 334-349.

Porter, L.W., "A Study of Perceived Need Satisfaction in Bottom and Middle Management Jobs," *Journal of Applied Psychology*, 1961, *45*, 1-10.

Young, G. C., "An Empirically Testable Model of Maslow's Theory of Human Motivation: Specification and Analysis," Unpublished Doctoral Dissertation, Stanford University, 1975.

VALUES OF AIR FORCE ACADEMY CADETS: A COMPARATIVE STUDY

Lawrence O. Short
Gerald L. Jones
Risdon J. Westen

The values of American military personnel are under close public scrutiny. Public concern involves the military services in general (Janowitz, 1971) and the service academies in particular (Lovell, 1964). This concern was shared by the authors of this study and led them to investigate two areas: The values of Air Foce Academy (AFA) cadets compared to their civilian counterparts and the values of fourth class AFA cadets (freshmen) compared to first class cadets (seniors).

For research purposes, values are defined as consisting of opinions about what is fair, just, or desirable. There appears to be a relationship between values and just about every aspect of our personal and professional lives: Values determine our frame of reference, our view of the world (Sikula, 1973).

Janowitz suggests (op. cit.) that the values of military and civilian management are becoming more similar and that differences are becoming less significant. His position is supported by Lovell (1964) who suggested that West Point graduates' values are affected by their military academy experience and that a process of professional socialization has begun. Lovell concludes that this process takes place in an environment which is closely related to and compatible with environments and with an educational emphasis that allows the developing of diverse values in preparation for increasingly diverse roles.

It is clear that value development and change is an important part of a college education. Indeed, Williamson and Biggs (1975) consider values so important that they list "perception of value options" as one aspect of total education, and DeCoster and Mable (1974) list "formation of a personal value system" as one of the five key developmental tasks of college students.

Clearly, the question of how students' values change is an important one. Brown (1972), for example, reports that some colleges have a significant effect on their students' values, attitudes, and eventual life-styles. After four years, students tend to become less authoritarian, less dogmatic, and less ethnocentric. They are more open, less conservative, somewhat less orthodox in their religious views, and more ready to express impulses. In addition, although changes are rather small, seniors occasionally seem to be more interested in intellectual issues than do freshmen, a finding generally consistent with McConnell (1972).

Based on the foregoing, in this preliminary study it was decided to investigate three areas: Area I - The logitundinal shift in values of entering AFA cadets; Area II - The values of senior AFA cadets compared to students in civilian colleges; and Area III - The shift in values of cadets during their four years at the AFA.

METHOD

Description of Instruments

The Polyphasic Values Inventory (PVI) by John T. Roscoe was used to measure values. It consists of twenty multiple-choice questions with responses that range along a continuum from *1 (most conservative) to 5 (most liberal). The questions cover selected* philosophical, political, economic, educational, social, personal-moral, and religious values. The questions and responses are intended to require the individual to make judgments of value rather than judgments of fact (Roscoe, 1965).

Procedure

To investigate Area I, data were collected from entering freshmen in the AFA Classes of 1973 through 1978 and examined for logitudinal shifts in values. Area II was investigated by comparing the values of a sample of 200 seniors from the AFA Class of 1974 with those of students in civilian colleges. Finally, Area III was investigated by comparing the freshmen and senior responses of the sample of 200 seniors from the AFA Class of 1974.

RESULTS

The following table shows the results of the longitudinal investigation of cadet responses to the PVI. Only item 2 showed any significant, sustained shift over the six-year period. Cadets entering later classes made a shift to a more liberal position regarding who should be allowed the right to vote.

The table also shows the modal responses for a normative sample (Roscoe, 1968) and pre/posttest modal responses for the Class of 1974. The differences on question 6 and 16 show the Academy samples to be slightly more conservative than the college student norm sample. It also can be seen that the Class of 1974 sample differed little on the pretest from the norm sample except for 6 and 16 where the cadets were slightly more conservative. The posttest, used to investigate differences in senior military and civilian college students, show the cadet sample differed from the college sample as follows:

1. Item 2 - liberal shift, more liberal than norm group
2. Item 3 - liberal shift, more liberal than norm group
3. Item 4 - liberal shift, more liberal than norm group
4. Item 5 - conservative shift, more conservative than norm group
5. Item 6 - conservative shift, more conservative than norm group
6. Item 17 - Liberal shift, more liberal than norm group

Area III was investigated using chi square to test the significance of shifts from the pretest to posttest for each PVI item. On item

TABLE 1
Modal Response Alternative by Question

	Class (By Year of Graduation)						College	CI 1974	
	1973	1974	1975	1976	1977	1978	Sample	Pre	Post
1. Nature of Science	2	2	2	2	2	2	2	2	2
2. Right to Vote	3	3	5	5	5	5	3	3	5
3. Communism	2	2	2	2	2	2	2	2	4&5
4. War	2	2	2	2	2	2	2	2	3
5. Foreign Policy	3	3	3	3	3	3	3	3	2
6. Private Enterprise	3	3	3	2	2	3	3	2&3	2
7. Labor Unions	2	2	2	2	2	2	2	2	2
8. Citizenship Education	3	3	3	3	3	3	3	3	3
9. Improvement in Schools	3	3	3	3	3	3	3	3	3
10 Academic Freedom	3	3	3	3	3	3	3	3	3
11. Equality of Man	3	3	3	3	3	3	3	3	3
12. Race Relations	5	5	5	5	5	5	5	5	5
13. Crime	2	2	2	2	2	2	2	2	2
14. Ethical Authority	3	3	3	3	3	3	3	3	3
15. Cheating on Tests	1	1	1	1	1	1	1	1	1
16. Sexual Relations	3	3	3	3	4	3	4	2	4
17. Alcoholic Beverages	3	3	3	3	3	3	3	3	3&4
18. Belief in God	2	2	2	2	2	2	2	2	2
19. Belief About the Bible	2	2	2	2	2	2	3	2	3
20. Responsibility of Man	1	1	1	1	1	1	1	1	1

1, although the modal response did not shift, the overall response shifted in the liberal direction ($x^2=22.45$, $P<.05$, $df=4$). On item 2 there was a strong and significant (x^2z30.74, $p<.05$, $df=4$) liberal shift. Almost half the posttest sample chose response alternative 5 as opposed to less than 30% of the pretest sample. A similar shift occured on item 3 ($x^2=66.10$, $p<.05$, $df=4$). About 60% of the pretest cases chose response alternatives 1, 2, and 3, while over 70% of the posttest cases chose response alternatives 4 and 5 (35.5%).

There was slight conservative to liberal shift on item 4 ($x^2=26.54$, $p<.05$, $df=4$) moving toward a moderate position on the issue of war. Responses clustered around alternatives 1 through 3, with the modal response shifting from response alternative 2 on the pretest to response alternative 3 on the posttest. Most seniors seemed to hold the value that we should not start a war but should be prepared to defend ourselves or our allies from attack.

Item 5, on foreign policy, showed a conservative shift toward nationalism and away from a position supporting world government ($x^2=21.70$, $p<.05$, $df=4$). This change, while in the reverse direction from previous shifts, is not surprising since the Air Force Academy is a military school with high emphasis on patriotism and love of country.

Item 6 did not show a significant shift ($x^2=6.10$, $p<.11$, $df=3$). On both the pre and posttest there was a generally moderate position supporting regulated free enterprise. There was also little shift on item 7 ($x^2=8.57$, $p<.07$, $df=4$). Cadets generally held a conservative position favoring right-to-work laws. There was, however, a slight tendency for those who were pro-union on the pretest to become more anti-union over the four years; but, generally, a position which began conservative remained so with no significant shifts.

On item 8, cadets started and ended with a moderate position that favored teachers guiding the experiences of children toward good citizenship ($x^2=5.99$, $p<.19$, $df=4$). Likewise on item 9, cadets started with a moderate position that subject matter in our nation's schools is generally adequate and maintained that position ($x^2=4.32$, $p<.36$, $df=4$).

Cadets started with a moderate position on academic freedom. No one on the pretest said it would be best if no controversial issues were discussed in class (the first response alternative). While the modal response remained alternative 3, there was significant shift ($x^2=35.01$, $p<.05$, $df=4$) to a more liberal position with almost 33% of the posttest sample advocating the greatest possible degree of academic freedom as opposed to 10% on the pretest.

Cadets stated a definite value on item 11: While people vary in their capacities, all persons are entitled to equal opportunity to develop those capacities. The response distribution on the posttest was virtually identical to the pretest, with the modal response being item 3 on both pretest and posttest ($x^2=4.78$, $p<.05$, $df=4$).

A conservative shift toward protecting society from criminals was noted on item 13 ($x^2=18.44$, $p<.05$, $df=3$). The frequency of response for alternative 2 (a dual program of reform and punishment is needed; the death penalty should be retained") increased from 48% of the pretest sample to 69% of the posttest sample, although this alternative remained the modal response for both groups. The most liberal response alternative was not selected by any testee in either the pretest or posttest group.

There was a significant ($x^2=10.16$, $p<.05$ $df=4$) convergence toward a moderate position on the value of ethical authority. Cadets seemed to feel that ethical issues are a function of society and that rules of conduct must be in accord with societal expectations.

As would be expected in an institution with a strong honor code, a vast majority of cadets on both pretest and posttest felt students are morally obligated to refrain from cheating and to discourage others from cheating. However, there was a slight liberal shift from pretest to posttest ($x^2=14.19$, $p<.05$, $df=3$), but no cadet in either the pretest or posttest group chose alternative 5 ("the student is entitled to take every opportunity to promote his own welfare").

The pretest to posttest difference on item 16 was significant ($x^2=95.31$, $p<.05$, $df=4$) and represented a strong conservative to liberal shift in attitudes toward restrictions on sexual intercourse. On the pretest, cadet response centered around 2, which specifies that sexual partners should be married but divorce should be permitted by society. On the posttest the modal response alternative shifted to alternative 4 ("much personal freedom is desirable, but the sex relationship is a personal one, and partners should be selected discretely").

The pretest to posttest shift for item 17 was significant ($x^2=22.09$, $p<.05$, $df=4$) and represents a slight conservative to liberal shift on values related to the use of alcoholic beverages. Cadets on the pretest held a moderate to liberal position (response alternative 3 = 53% of total; response alternative 4 = 26% of total). On the posttest, the two alternatives were again most frequently selected, but by almost equal percents of the sample.

Data on item 18 about belief in God showed no significant shift from pretest to posttest ($x^2=5.46$, $p<.20$, $df=4$). Cadets on both the pretest and the posttest most frequently selected response alternative 2, acknowledging the existence of a personal God common to all faiths. Similarly, item 19 on belief about the Bible showed no significant shift from pretest to posttest ($x^2=4.82$, $p<.31$, $df=4$).

As with the previous two items, no significant pretest to posttest shifts were noted on the item concerned with responsibility of man ($x^2=5.65$, $p<.23$, $df=4$). Cadets took a very conservative position on the pretest and essentially maintained this position although there was a slight trend toward a more moderate position.

One other finding seems worth noting. In comparing pretest and posttest responses of members of AFA Class of 1974, it became obvious that, except for two items, there was no significant increase in the homogeneity of cadet responses to the values statements. Entering cadets showed a diversity of responses, and four years at the Academy did not seem to eliminate any of the diversity.

DISCUSSION AND CONCLUSION

This preliminary examination of the data suggests that the values of entering cadets did not greatly differ from year to year over the period covered. It also suggests that the values of AFA cadets are similar to those of their civilian college contemporaries and that they become somewhat more liberal as seniors. These findings are in consonance with those reported by Brown for civilian college students. The liberal shift may, however, be a function of changing societal values rather than a result of the Academy's educational processes and programs.

Since numerous value changes occurred while a diversity of opinions remained and our results were similar to civilian student data, it is clear that cadets do not become young automatons who look, act, and think alike in conservative "military" manner. At the same time, as a group they espouse the values of patriotism, individual honesty, and responsibility.

This preliminary exploration of cadet values and changes in such values raises many questions about how and why changes occur and the influence of societal and environmental factors on values. More current norms and much more carefully controlled studies are clearly necessary.

REFERENCES

Brown, R. D., *Student development in tomorrow's higher education: A return to the academy*. Washington, D.C.: ACPA Student Personnel Series No. 16, 1972.

DeCoster, D. A. & Mable, P., (Eds.) *Student development and education in college residence halls*. Washington, D.C.: ACPA Press, 1974.

Janowitz, M., *The professional soldier*. New York: Macmillan, 1971.

Lovell, J. P., The professional socialization of the West Point cadet. In Janowitz, M. (Ed.). *The new military: Changing patterns of organization*. New York: Wiley & Sons, 1964.

McConnell, T. R., Do colleges affect student values? *Change*, 1972, *4*, 9, 63-64.

Roscoe, J.T. Report of first research with the Polyphasic Values Inventory. *Journal of Research Services, Colorado State College*, 1965, *5*, 3-12.

Roscoe, J. T., Ritter, C. E., Teglovic, S. & Thayer, J. D., American college student values: Preliminary report of a nationwide survey. *Colorado Journal of Educational Research*, 1968, *8*, 3-26.

INFLUENCING BUREAUCRATIC MACROSYSTEMS: THEORY AND STRATEGY FOR ORGANIZATIONAL CHANGE

by
Raymond L. Forbes Jr., Ph.D.
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

Many theorists have long held that bureaucratic organizational forms are particularly difficult targets for change. This resistance to change has been variously attributed to size (Sayles, and Chandler, 1971), emphasis on formal rules and procedures (Zaltman and Duncan, 1977), and investment in maintenance of the status quo (Swingle, 1977). Bureaucratic macrosystems or suprasystems are giant organizations composed of other large organizations (e.g., General Motors, the University of California, and the Department of Defense). Compared to smaller organizations macrosystems appear to be much more difficult to influence because of their progressively higher levels of complexity and their sheer physical and administrative bulk. These suprasystems wield immense power in our society as well as contributing to its major problems. For many persons intent on changing them in discernible ways bureaucratic macrosystems have become the fabled immovable object.

To date, much of the work of the professional organizational change practitioner has been implicitly grounded in the ideas of social psychologist, Kurt Lewin. In particular, Lewin's formulation of social change as a three-phased process has found wide acceptance among those seeking a rational explanation for change. The notion of change as a continuing sequence of unfreezing, changing, and refreezing steps has provided a simple and practical concept for visualizing the change process. However, the application of Lewin's theory at the macro-bureaucratic level of analysis has provided useful but, thus far, incomplete understanding of how large-scale change processes actually occur.

Essentially, Lewin's theory can describe *what* needs to happen in an organizational change program but not answer *why* it needs to do it nor *how* to go about accomplishing it. The potential macrosystem change agent must seek his own ways to create organizational disequilibrium, introduce a different way of functioning, and then restabilize the organizational system so that it will successfully maintain its altered state of operating. When you're trying to change, for example, the way the U.S. State Department does business effecting even the smallest noticeable change can be tremendously time and energy consuming (Marrow, 1974).

In looking at how macrosystems change, the experience of the U.S. military may provide some significant insights. Three of the military services (the Army, the Navy, and the Air Force) have been extensively experimenting with large scale organizational planned change efforts for the past seven years. Drawing on this practical experience from military bureaucratic macrosystems, two alternate theories for looking at large system change have developed. Like McGregor's X-Y dichotomy of managerial behavior (McGregor, 1960) each theory is based upon a contrasting set of basic assumptions. These assumptions are rooted in people's ideas about how organizations appear to them to work and change.

One theory is based upon a mechanical or "clockworks" idea of organizational functioning. The alternate theory is founded on a contrasting notion of an organization as an organic or living system. Each, on the basis of its assumptions, has a corresponding set of strategies designed to produce organizational change. The strategies, in turn, provide the direction and guidelines for the tactics required to implement a change program.

The first approach to organizational change, "critical mass" theory, states that significant change will occur in a large orgnizational system when the sum of individual incremental changes in its components reaches some critical value. This method is primarily evolutionary in nature in that it depends on the aggregation of similarly disposed people in order to produce change. This critical mass approach has its analog in nuclear physics. Significant change (a nuclear chain reaction) occurs in fissionable material only when the total radiation level, as additional material is added, reaches some crucial point (criticality). The assumptions which underlie this theory are:

1. The complete organizational system is far too complex to operate upon as an entity.
2. An autonomous large system is a myth; the reality is that only individuals and small collectives of persons comprise the actual organization.
3. Complicated large system functioning can only be really affected by the initiation of an expanding process of influence through its component individual and group members.
4. If enough individuals and groups can be persuaded to change then eventually the momentum that is built up will reach a critical level and result in system movement.
5. The change process must permeate a significant portion of entire organizational structure if the desired changes are to be realized and sustained.

An alternative theory of macro-bureaucratic change tentatively called "realignment theory" postulates that significant premeditated change will occur in a large system when the constellation of external and internal forces acting upon it are deliberately realigned in some new configuration. This means of change induction is principally revolutionary in nature since it tends to produce change rapidly over a broad area. A rough biological analogy for this theory is the concept of "homeostasis" or the adaptive adjustment of a living organism when confronted with a potentially disruptive force in its environment. The central assumptions for this theory are:

1. The entire organizational system of component parts can be conceived of and operated upon as if a single entity.
2. The organizational system has an existence of its own (synergy) that is greater than the sum of its component individuals and groups.
3. The system is amenable to a logic-based rational analysis of the forces that hold it in balance.
4. Manipulation of variables affecting system balance can be managed from within or without the organizational structure.
5. No change will occur in a large system until some unknown level of bureaucratic inertial or resistance is overcome.

Strategies for implementing organizational change flow directly from each of the two basic sets of assumptions. Many are conceived so as to employ "organizational jujitsu" or the use of internal leverage points to redirect the thrust of change pressures. This idea applies the martial art philosophy of capitalizing on an opponent's strengths or using the very size, complexity, and diversity of resources of the bureaucratic macro-system as a means for moving it to self-change. For example, this might take the form of blending the upward mobility ambitions of key subsystem managers to induce either collaborative or competitive financial outcomes.

If it is assumed that the principal means for producing significant organization change is through incremental aggregation then a family of supportive change strategies can be developed. A few examples of these strategies might be:

- The "*Cue Ball*" method - An organizational subunit is identified for change, invested with a special mandate for change by top management and is directed to produce change in each of the other units of the organization. Individual units that change in the desired manner as a consequence of the interaction continue to receive organizational support. The initiators of change continue to operate in this manner until a majority of units function in the desired manner. *The non-adopters become organizational casualties.*
- "The "*Quick Opener*" method - An unproductive unit is selected by top management to use as an example of the change it desires, a plan for major alteration of the unit is conceived without its knowledge and a surprise and massive demand for change is made upon its leadership by their superiors. When sufficient volume and intensity of change pressure has been generated the target unit leadership either changes to avoid being damaged, develops successful counter-measures or leaves. Other units observing the process react in the desired manner to protect themselves from the same fate. *Observational units may get the wrong change method.*
- The *Demolition Derby*" method - A currently stable but potentially powerful organizational subsystem is chosen to be the target; a smaller but energetic subsystem is set-up to forcibly compete with it for resources to achieve the wanted change on a win-lose basis. A massive explosion ensues as a result of the competition, destroying or seriously damaging the smaller unit's effectiveness. A large amount of energy is released into the system from the competition in a chain reaction that rapidly affects other parts of the system. *The results are often catastrophic, in unpredictable directions, with some destructive residual long-term fallout.*
- The "*Reactor Pile*" method - This strategy is a variation of the "Demolition Derby" approach wherein unstable parts of the organization are administratively mandated to cooperate in a change program. When sufficient numbers are in place, enough energy begins to be released to sustain a low level chain reaction of change; moderator devices (managers) are placed throughout the system to keep the reaction under control and produce a new level of power output from the whole system. *Sometimes the unit leadership involved in this approach "burns out," is reduced to an inert but stable condition, and may require replacement.*

For the proponents of the realignment theory of organizational change, strategic options might include:

- The "*War Gaming*" method - A dynamic statistical model of the organizational system is developed using computer based techniques. Critical variables are experimentally manipulated within various scenarios, contingency plans are developed, and the plans are implemented when certain "trigger" events occur. The result is often large-scale movement of people, fiscal, and other resources from one location to another in a carefully planned manner. *The system often returns to its previous state shortly after the disturbance is over; little actual change occurs.*

- The "*Pressure Cooker*" method - Various new ingredients (change ideas) are added by outsiders from top management, the pressure builds up, and the contents of the "pot" become inextricably fused together. *The old organizational system is often unrecognizable in the new "stew".*

- The "*Crock Pot*" method - This is a variation of the "Pressure Cooker" theme that employs less containment, lower heat application, and little build-up of pressure. It requires a longer period for mixing and slow alteration of the structure of the ingredients and produces less dramatic but equally as palatable results as the more forceful approach. *Some units may come apart under the pressure build-up.*

- The "*Swine Flu*" method - A sufficient number of negative organizational symptons occur for the top leadership to recognize wide scale discomfort. A group of experts is formed to develop a planned change vaccine to provide an acceptable degree of relief. The organizational vaccine is produced and tested under laboratory conditions in a test subsystem, and mass innoculations are administered throughout the organization. *The immunization works for the majority but produces some unanticipated and undesirable side effects (like alienation and counter reactions).*

The evidence, to date, for a workable and comprehensive theory of macro-system change is scanty at best. The great majority of change efforts have concentrated upon much smaller systems. At our current stage of macro-system change development it remains largely a moot point whether or not large scale, consciously planned efforts can be effective. Sufficient indications do seem to exist, however, to show that an increasing number of such efforts are being attempted. It is hoped that for those considering change in a bureaucratic macrosystem this paper may serve as a point of departure.

REFERENCES

Marrow, A., *The Practical Theorist*. New York: Basic Books, 1969.

McGregor, D., *The Human Side of Enterprise*. New York: McGraw-Hill, 1960.

Sayles, L. R. and M.K. Chandler, *Managing Large Systems: Organizations of the Future*. New York: Harper & Row, 1971.

Swingle, P. G., *The Management of Power*. La Jolla, CA: NTL/Learning Resources Corp., 1976.

Zaltman, G. and R. Duncan. *Strategies for Planned Change*. New York: John Wiley, 1977.

AUTHOR INDEX